Theory and Application of a Pólya Urn with Non-Linear Feedback

Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.

eingereicht an der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
der Universität Augsburg

von
Thomas Michael Gottfried

Augsburg, 22. Juli 2024

Universität
Augsburg
University

Betreuer und Erstgutachter:	Prof. Dr. Stefan Großkinsky
Zweitgutachter:	Prof. Dr. Markus Heydenreich
Drittgutachter:	Prof. Dr. Enrico Scalas
Tag der mündlichen Prüfung:	25. Oktober 2024

Abstract

Generalized Pólya urns with non-linear feedback are an established probabilistic model to describe the dynamics of growth processes with reinforcement. Depending on the feedback function, this process exhibits monopoly, where a single, random agent (or colour) achieves full market share, or it converges to a deterministic limit point. This work provides a comprehensive account of the properties of this process for fairly general feedback. In the monopoly case, we derive results on the prediction of the monopolist for large initial market size and in the deterministic case, we give an asymptotic description of the evolution of market shares. We describe in detail how monopolies emerge in a transition from sub-linear to super-linear feedback via hierarchical states close to linearity. Moreover, we derive a scaling limit for the dynamics and characterize the fluctuations in a functional central limit theorem. By choosing a different approach, we generalize known results on the number of steps won by the losing agents in the case of strong monopoly, which even provides information on dependencies between several losers and on the tail distribution of further related quantities like the time of monopoly. As a seeming paradox, losers with feedback close to the identity are most likely to win in many steps. Finally, we suggest an extended version of the non-linear Pólya urn as a model for the dynamics of wealth distribution within an economy, where we particularly highlight the empirical observation that wealth is distributed significantly more unequal than wages. This allows some interesting predictions for future developments.

Acknowledgements

First of all, I would like to thank my supervisor Prof. Stefan Großkinsky for suggesting this exciting topic and for his extensive support and inspiration. I particularly appreciate that he gave me the freedom to pursue my own research interests. I greatly enjoyed our frequent mathematical and non-mathematical talks. Second, I would like to express my gratitude to Prof. Lothar Heinrich, who introduced me to the thrilling field of stochastics and extended my historical knowledge in numerous long discussions. I am grateful to Dr. Günter Hinrichs for convincing me to start this project and to Claudia Stucke for always spreading good mood at work. I would also like to thank my long-standing friend Fabian Endres for enriching my time in Augsburg (and before) with much fun and motivation. Moreover, I want to pronounce a heartfelt "merci beaucoup" to my girlfriend for linguistic assistance and emotional encouragement. Finally, I express my deep gratitude to my parents for their unconditional love and their dedicated support for whatever I am doing.

Contents

Notation directory

$[\Omega, \mathcal{A}, \mathbb{P}]$	probability space
$\mathbb{E}$	expectation with respect to $\mathbb{P}$
Var	variance with respect to $\mathbb{P}$
$\mathbb{N}$	$:= \{1, 2, \ldots\}$ natural numbers
$\mathbb{R}$	real numbers
$[A]$	$:= \{1, \ldots, A\}$ for $A \in \mathbb{N}$
Δ_{A-1}	$:= \{(x_1, \ldots, x_A) \in [0,1]^A \colon x_1 + ... + x_A = 1\}$ $(A-1)$-dimensional simplex
$T\Delta_{A-1}$	$:= \{(x_1, \ldots, x_A) \in \mathbb{R}^A \colon x_1 + ... + x_A = 0\}$ tangent space
Δ^o_{A-1}	$:= \{(x_1, \ldots, x_A) \in (0,1)^A \colon x_1 + ... + x_A = 1\}$ open simplex
B^o	interior of a Borel set B
$\overline{B}$	closure of a Borel set B
$\#S$	cardinality of a set S
S^c	complement of a set S
$\delta_{i,j}$	Kronecker delta
$e^{(i)}$	$:= (\delta_{i,j})_{j=1,\ldots,A} \in \mathbb{R}^A$ i-th unit vector
$x_n \sim y_n$	$:\Leftrightarrow \lim_{n\to\infty} \frac{x_n}{y_n} = 1$ for real sequences (x_n), (y_n)
$x_n \prec y_n$	$:\Leftrightarrow \limsup_{n\to\infty} \frac{x_n}{y_n} < \infty$ for real sequences (x_n), (y_n)
$x_n \asymp y_n$	$:\Leftrightarrow x_n \prec y_n$ and $y_n \prec x_n$ for real sequences (x_n), (y_n)
o, O	Bachmann-Landau notation
$\|\cdot\|$	Euclidean norm
$\|\cdot\|_\infty$	supremum norm
$\langle \cdot\,, \cdot \rangle$	Euclidean scalar product
$\mathbb{D}$	Skorokhod space, i.e. space of càdlàg functions
CDF	cumulative distribution function
ODE	ordinary differential equation
SDE	stochastic differential equation
RODE	random ordinary differential equation
LLN	law of large numbers
CLT	central limit theorem
w.l.o.g.	without loss of generality
Φ	CDF of the normal distribution
log	natural logarithm
$\mathbb{1}$	indicator function

For to every one who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away.

- Gospel of Matthew, 25:29, RSV

Chapter 1

Introduction

The evolution of inequality and its determinants is a much discussed issue in research and public debates, not least due to the enormous public impact of Thomas Piketty's work [110, 109]. As early as 1897, Vilfredo Pareto observed that wealth follows a power-law distribution, i.e. the share of population with wealth exceeding $w > 0$ is approximately given by $cw^{-\alpha}$ for a constant $c > 0$ and an exponent $\alpha > 0$. Moreover, he predicts that α is fairly stable in time and for different places. Since then, this stunning observation has been confirmed by numerous studies (e.g. [125, 19]) and was later promoted by the Nobel prize winning economist Paul A. Samuelson [116] as "Pareto's law". For Germany and the USA, Figure 1.1 shows the distribution of net personal wealth per adult, which is defined as the total value of non-financial and financial assets (housing, land, deposits, bonds, equities, etc.) held by individuals, minus their debts. Figure 1.1 is consistent with Pareto's law.

Unveiling the mechanism behind the dynamics of wealth by the means of mathematics can support political and ethical debates on how to achieve a more equitable world. Simple models assuming independent return rates (like the Black-Scholes model) cannot account for these heavy-tailed wealth distributions, so more refined research is necessary. Opposing classical economic theory, the famous economist W. Brian Arthur [5] identifies increasing return rates as a main driver in the evolution of any free market, i.e. the return rates on capital depend on the amount of capital a person or household owns. In simple terms, the richer you are, the faster your wealth grows. This phenomenon is sometimes referred to as Matthew effect, which traces back to the introductory quotation from the Gospel of Matthew. Diverse explanations for increasing returns can be conceived, like lower risk aversion of the rich due to higher risk-bearing potential. Empirical studies in [53] confirm this idea, but on the other hand increasing returns can even be found within similar asset groups. Hence, risk does not exclusively drive increasing returns and other factors like higher skills, political influence, informational advantages, decreasing costs of debt, tax-evasion or lower transaction costs are relevant, too. Moreover, some asset classes are barely available for ordinary people, like private equity fonds, art or other value increasing luxury goods. This work will focus on the impact of increasing returns on wealth inequality rather than on the underlying reasons. For a mathematical treatment of this phenomenon, Arthur [5] suggests a generalization of Pólya's urn model, on which this work provides some new, comprehensive results.

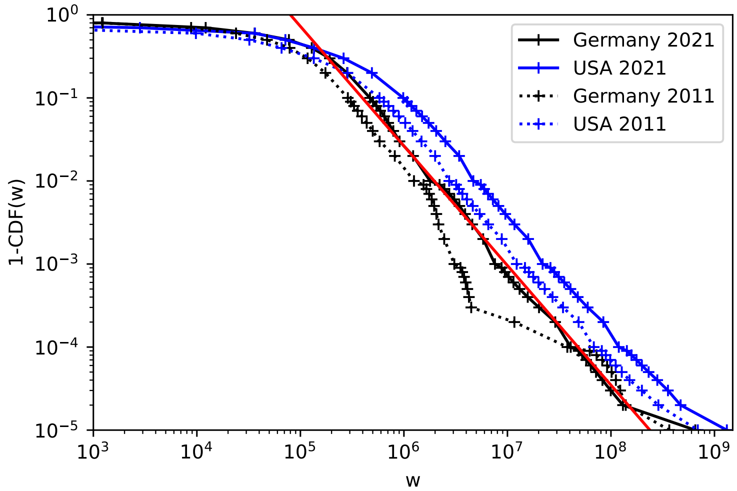


Figure 1.1: $1 - CDF$ (Cumulative Distribution Function) of net personal wealth (purchasing power parity, equal split adults) in Germany and the USA in 2011 and 2021 in Euro resp. US-Dollar according to [111]. Least square fit (red line) estimates a Pareto exponent of 1.44 for Germany 2021.

In general, Pólya's urn is a fundamental model for growth processes. Given an urn containing balls of several colours, one ball is drawn uniformly at random and is put back to the urn together with another ball of the same colour. Then, the procedure is repeated. Due to its simplicity, this model allows for a rigorous mathematical analysis, but it is still flexible enough to find applications in numerous fields like economics, psychology, evolutionary game theory, statistics or image reconstruction (see Pemantle [107] and references therein for an overview). More recently, [122, 114] use Pólya urns in the context of cryptocurrencies.

The model was proposed in 1923 by Pólya and Eggenberger [52] to describe the spread of infectious disease, but it has been considered by Markov [94] before. Since then, many generalizations of their idea were studied. For example in 1980, Hill, Lane and Sudderth [73] extended the original model to non-linear feedback, which was later promoted by Arthur [5] as a fundamental model for reinforced competition, e.g. between nations, companies, technologies or individuals. Another frequently studied extension of Pólya's urn model is Friedman's urn [63] from 1949, where in each step depending on the drawn colour a certain set of balls is added (or even removed) from the urn. Moreover, reinforced random walks as initiated by Coppersmith and Diaconis [42] in 1986 are closely related to Pólya's urn [45, 123, 44], as the probability to walk along an edge depends on how often it was crossed before. Aletti and Crimaldi [3, 2, 4] include a rescaling mechanism to inhibit long-range dynamic dependencies and [75, 21, 14, 41] deal with innovation in Pólya's model. In addition to that, Pólya urns with infinitely many colours were introduced in [10, 93, 24] and have been addressed in recent publications [9, 80, 117]. Further generalizations of Pólya's urn to more complex replacement schemes (like several interacting urns) have been studied in [96, 43, 18, 86, 115, 121, 119]. A comprehensive overview on different urn models is provided in [92, 88, 82].

This work focuses on Pólya urns with non-linear feedback as proposed in [5, 73, 30]. Let us first illustrate the idea of the model by means of an example. In the near future, customers who intend to buy a new car will have the choice between several different technologies like modern cars powered by fossil or synthetic fuels, hydrogen or batteries. Although electric cars seem to be in the pole position in the race for the future car market, it is still open which technology will win or whether there will be a mixture of different technologies. The decision which technology a customer chooses basically depends on three factors. First, each technology has an intrinsic deterministic attractiveness or fitness, which is e.g. the reason why cars once replaced horses. Second, the decision depends on the choice of earlier customers. For example, if many have bought an electric car before, there will be a dense charging infrastructure and thus electric cars get more attractive for future customers. A second argument for this reinforcement is that high revenues in the past provide financial means for a faster technological development as well as cheaper prices because of lower production costs per unit. The resulting overall attractiveness of technology i is now modeled by a hypothetical feedback function $F_i(X_i) \geq 0$ depending on the number $X_i \in \mathbb{N} = \{1, 2, 3 \ldots\}$ of customers who have chosen technology i before. High values of $F_i(X_i)$ indicate high attractiveness of technology i. A typical example is $F_i(k) = \alpha_i k^\beta$, where $\alpha_i > 0$ models the intrinsic attractiveness and $\beta > 0$ the reinforcement effects in the market. The third determinant of customers' decision is their personal preference, which is difficult to include in a deterministic model, so probabilistic approaches are more appropriate. We assume that customers enter the market sequentially and have full information. Given the current state $(X_1, \ldots, X_A)$ of the market, a customer will opt for technology i with probability

$$\frac{F_i(X_i)}{F_1(X_1) + ... + F_A(X_A)},$$

where $A \geq 2$ is the number of different technologies. The market size $X_1 + \ldots + X_A$ increases by one in each step. If $F_i(k) = k$, then this corresponds to the original Pólya urn [52]. Depending on the feedback function, monopoly may occur, where one technology achieves full market share, as well as random or deterministic non-zero asymptotic market shares for several technologies. The monopolist is in general random and depends on the behaviour of the young market. Analyzing which feedback function leads to which properties provides an understanding of the determinants of the behaviour of markets. For example, a natural question of interest is to which extent the fitness α_i and the initial market shares affect the long-time limits.

Mathematically, this setup corresponds to a discrete-time Markov process, which is called a (generalized) non-linear Pólya urn in the following. Properties of these urns have been examined before, often focused on polynomial feedback functions [73, 83, 49, 84, 103, 81, 40, 97] or symmetric models with $F_i \equiv F$ [102, 101, 98]. In applications, feedback functions are usually a hypothetical construction, which can barely be measured in real systems (like in [60, 124]) similar to utility functions in economic situations. Thus, a general mathematical understanding without restrictive conditions on F_i is important. This work investigates the behaviour of non-linear Pólya urns for a very general class of feedback functions, which unveils some new exciting facets of this process, like random weak monopoly (Section 3.3) or the "loser paradox" (Section 5.2). An important restriction is, however, that F_i depends only on X_i, which excludes stationary limit cycles as studied e.g. in [43].

This thesis is structured as follows: After formally introducing the model in Chapter 2, we study the long-time behaviour of the process in detail in Chapter 3. In Chapter 4, we derive a scaling limit for the whole dynamics, including an asymptotic description of fluctuations. Next, in the case of monopoly, we generalize known results on the number of steps (or customers) won by the losing agents (or technologies) in Chapter 5. Finally, in Chapter 6, we give an example of how this urn model can be used for applied purposes, in particular as a model for the distribution of wealth. Section 2.3 provides a more detailed overview on the main contributions of this work. Most results of this work have been published with Stefan Großkinsky in [67, 69, 68].

Chapter 2

Preliminary Definitions and Results

This chapter formally introduces the non-linear Pólya urn model. We establish important notation and terminology used throughout this work, including the fundamental concept of exponential embedding (Section 2.1). In Section 2.2, we present a selection of important results on this model, which are related to our work. Finally in Section 2.3, we outline the structure and main novelties of this work.

2.1 The model and the exponential embedding

We now formally define the model. All random variables are defined on some large enough probability space $[\Omega, \mathcal{A}, \mathbb{P}]$. Let $A \geq 2$ be the number of agents (corresponding to technologies in the introduction) and $F_i \colon \mathbb{N} \to (0, \infty)$ the feedback function of agent $i \in [A] \coloneqq \{1, \ldots, A\}$. We define a homogeneous, discrete-time Markov process $(X(n))_{n\in\mathbb{N}_0} = ((X_1(n), \ldots, X_A(n))_{n\in\mathbb{N}_0}$ on the state space $\mathbb{N}^A$ with initial condition $X(0) = (X_1(0), \ldots, X_A(0)) \in \mathbb{N}^A$ and transition probabilities

$$\mathbb{P}\left(X(n+1) = X(n) + e^{(i)} \,\middle|\, X(n)\right) = \frac{F_i(X_i(n))}{F_1(X_1(n)) + \ldots + F_A(X_A(n))},\ i \in [A]\,, \tag{2.1}$$

where $e^{(i)} = (\delta_{i,j})_{j=1}^A$ is the i-th unit vector. We denote by $N \coloneqq X_1(0) + \ldots + X_A(0) \geq A$ the initial market size. Whenever needed, we set $F_i(0) = 0$, and whenever useful, we take smooth extensions $F_i : (0, \infty) \to (0, \infty)$ to the positive real line.

We interpret $X_i(n)$ as the number of customers of agent i at time n and define the corresponding time-inhomogeneous Markov process $(\chi(n))_{n\in\mathbb{N}_0}$ of **(market) shares**

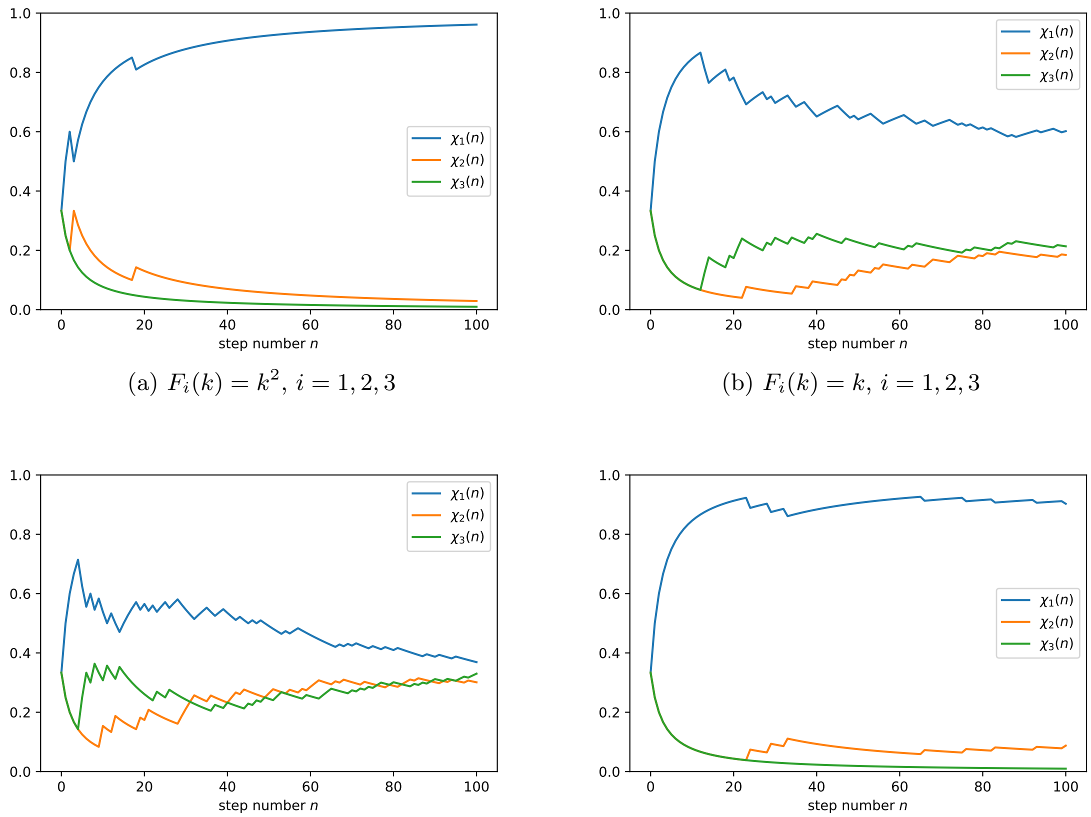
$$\chi_i(n) \coloneqq \frac{X_i(n)}{N+n} \in (0,1),\ i = 1, \ldots, A,\ n \in \mathbb{N}_0\ ,$$

with $\chi(n) = (\chi_1(n), \ldots, \chi_A(n)) \in \Delta^o_{A-1}$, where

$$\Delta^o_{A-1} \coloneqq \{(x_1, \ldots, x_A) \in (0,1)^A \colon x_1 + \ldots + x_A = 1\}$$

is the interior of the unit simplex $\Delta_{A-1} \coloneqq \{(x_1, \ldots, x_A) \in [0,1]^A \colon x_1 + \ldots + x_A = 1\}$. Moreover, we establish the notation

$$\chi(\infty) \coloneqq \lim_{n\to\infty} \chi(n)$$



(a) $F_i(k) = k^2$, $i = 1, 2, 3$

(b) $F_i(k) = k$, $i = 1, 2, 3$

(c) $F_i(k) = \sqrt{k}$, $i = 1, 2, 3$

(d) $F_i(k) = k \log(k+1)$, $i = 1, 2, 3$

Figure 2.1: Simulated evolution of the market shares for the first 100 steps of a non-linear Pólya urn with different feedback functions. Here $A = 3$ and $X(0) = (1, 1, 1)$.

for the long-time market share whenever it exists. We will see throughout this work that $\chi(\infty)$ is well defined in all generic situations, but it is possible to construct counterexamples (see Example 3.28). For later use, we introduce the notation

$$p(k, x) = (p_i(k, x))_{i \in [A]} = \left(\frac{F_i(kx_i)}{F_1(kx_1) + \ldots + F_A(kx_A)} \right)_{i \in [A]} \tag{2.2}$$

for the transition probabilities, where $k \in \mathbb{N}$ and $x = (x_1, \ldots, x_A) \in \Delta_{A-1}$. Figure 2.1 shows four simulations of this process for different feedback functions, where the different regimes (outlined in detail in Section 3.4) are visible.

A useful alternative construction of the process is provided by the so-called **exponential embedding**. We take independent random variables $\tau_i(k)$, $i \in [A]$, $k \in \mathbb{N}$, where $\tau_i(k)$ is exponentially distributed with rate parameter $F_i(k)$. For each i we define the corresponding continuous-time counting process $\left(\Xi_i(t)\right)_{t\geq 0}$ with

$$\Xi_i(t) = \Xi_i^{(X_i(0))}(t) := \max\left\{ l \in \mathbb{N}_0 : \sum_{k=0}^{l-1} \tau_i(X_i(0)+k) \leq t \right\} + X_i(0),\ t \geq 0. \tag{2.3}$$

These are independent birth processes with $\Xi_i(0) = X_i(0)$, where the time between the k-th and $(k+1)$-th event of Ξ_i is given by $\tau_i(X_i(0)+k)$. If $0 = t_0 < t_1 < t_2 < ...$ is the sequence of jump-times of the process $\Xi(t) = (\Xi_1(t), \ldots, \Xi_A(t))$, i.e.

$$t_{n+1} = \min\left\{ t > t_n : \Xi(t) \neq \Xi(t_n) \right\},$$

then **Rubin's theorem** (published in [45]) states that the jump chain $(\Xi(t_n))_{n\in\mathbb{N}_0}$ has the same distribution as the process $(X(n))_{n\in\mathbb{N}_0}$. Thus we can define:

$$X(n) := \Xi(t_n) \tag{2.4}$$

The proof of Rubin's Theorem is based on the lack of memory property of exponentials, i.e. given $\Xi(t_n) = (x_1, \ldots, x_A)$ we have

$$\begin{aligned} \mathbb{P}\left(\Xi(t_{n+1}) = \Xi(t_n) + e^{(i)} \,|\, \Xi(t_n) = (x_1, \ldots, x_A)\right) &= \mathbb{P}\left(\tau_i(x_i) = \min_{j\in[A]} \tau_j(x_j)\right) \\ &= \frac{F_i(x_i)}{F_1(x_1) + \ldots F_A(x_A)}. \end{aligned}$$

In fact, the birth processes $\Xi_i(t)$ can explode as the sum $\sum_{k=X_i(0)}^{\infty} \tau_i(k)$ might be finite. We therefore define the random **explosion times**

$$T_i(X_i(0)) := \sum_{k=X_i(0)}^{\infty} \tau_i(k) \in (0,\infty],\ i \in [A]. \tag{2.5}$$

In the following we are especially interested in the occurrence of monopoly, which requires some definitions.

Definition 2.1. For $i \in [A]$, we define the events

1. **weak monopoly**

$$wMon_i(\chi(0), N) := \left\{ \omega \in \Omega : \lim_{n\to\infty} \chi_i(n)(\omega) = 1 \right\} = \left\{ \lim_{n\to\infty} \chi_i(n) = 1 \right\},$$

i.e. the market share of agent i converges to one;

2. **strong monopoly**

$$sMon_i(\chi(0), N) := \left\{ \lim_{n\to\infty} \sum_{j\neq i} X_j(n) < \infty \right\},$$

i.e. agent i wins in all but finitely many steps;

3. **total monopoly**

$$tMon_i(\chi(0), N) := \Big\{\forall n \geq 0 \, \forall j \in [A] \setminus \{i\} : X_j(n) = X_j(0)\Big\},$$

i.e. agent i wins in all steps.

On the event $wMon_i(\chi(0), N)$, we call agent i the **monopolist** or **winner** and all other agents $j \neq i$ **losers**.

Obviously, a total monopoly is also a strong monopoly and a strong monopoly always implies a weak monopoly. Via exponential embedding one can express the event $sMon_i(\chi(0), N)$ by the explosion times through

$$sMon_i(\chi(0), N) = \bigcap_{j \neq i} \Big\{T_i(X_i(0)) < T_j(X_j(0))\Big\} \tag{2.6}$$

as equality of finite explosion times has probability zero. With the observation

$$T_i(X_i(0)) < \infty \Leftrightarrow \mathbb{E}T_i(X_i(0)) = \sum_{k=X_i(0)}^{\infty} \frac{1}{F_i(k)} < \infty \Leftrightarrow \sum_{k=1}^{\infty} \frac{1}{F_i(k)} < \infty,$$

(proven by Feller in [54] and sometimes referred to as Feller-Lundberg-Criterion) one can easily derive the following generally known criterion for the occurrence of strong monopoly from [45].

Theorem 2.2. *[45, Appendix] Strong monopoly occurs with probability one, i.e.*

$$\mathbb{P}\left(\bigcup_{i=1}^{A} sMon_i(\chi(0), N)\right\} = 1,$$

if and only if

$$\sum_{k=1}^{\infty} \frac{1}{F_i(k)} < \infty \tag{M}$$

for at least one $i \in [A]$, otherwise the probability is zero.

If (M) holds, the density of the explosion time $T_i(X_i(0))$ (see Lemma 5.2) as a sum of exponential variables has support on the whole positive real line for all choices of F_i. So the probability of $sMon_i(\chi(0), N)$ is positive if and only if agent i fulfills (M) and the monopolist is random among all agents $i \in [A]$ that satisfy (M). For the polynomial case $F_i(k) = \alpha_i k^\beta$, $\alpha_i > 0$, $\beta \in \mathbb{R}$, $i \in [A]$, Theorem 2.2 implies that strong monopoly occurs if and only if $\beta > 1$.

On the other hand, when no agent fulfills (M), $X_i(n) \to \infty$ almost surely for all $i \in [A]$ and we have the following consistency property.

Proposition 2.3. *Assume that none of the F_i satisfies (M). Define a 'partial' Pólya urn process $\tilde{X}(n)$ for a subset $B \subset [A]$ of agents with the same feedback functions F_i and initial condition $\tilde{X}(0) = (X_i(0) : i \in B)$. Then the process $\big(\tilde{X}(n)\big)_{n\in\mathbb{N}_0}$ can be identified as a (random) subsequence of $\big(X_i(n) : i \in B\big)_{n\in\mathbb{N}_0}$.*

Proof. The independence property of the exponential embedding provides a canonical coupling of the processes $\tilde{X}$ and X. For that, define recursively $s_0 = 0$ and

$$s_{n+1} := \inf\{s > s_n \colon \exists i \in B \colon \Xi_i(s) \neq \Xi_i(s_n)\}\,.$$

Note that $s_n < \infty$ is well defined for all $n \geq 0$, since none of the F_i fulfill (M). Then set $\tilde{X}_i(n) := \Xi_i(s_n)$ for $i \in B$, which directly implies the claim since (s_n) is a subsequence of (t_n). □

In particular, if one of the limits

$$\tilde{\chi}(\infty) := \lim_{n\to\infty}\left(\frac{\tilde{X}_i(n)}{\sum_{j\in B}\tilde{X}_j(n)}\right)_{i\in B}\,,\quad \chi^B(\infty) := \lim_{n\to\infty}\left(\frac{X_i(n)}{\sum_{j\in B}X_j(n)}\right)_{i\in B}$$

exists, then so does the other and both have the same distribution. This implies further neutrality of the limit $\chi(\infty)$ in the sense of [78], so that it has a (possibly degenerate) Dirichlet distribution on Δ_{A-1}, whenever it exists. In the degenerate case, the Dirichlet distribution is either deterministic or concentrated on the boundary of Δ_{A-1}, e.g. when $\mathbb{P}\left(\bigcup_{i=1}^A wMon_i(\chi(0), N)\right) = 1$. This will be discussed in several examples in Sections 3.2 and 3.3. Note that in the case of weak monopoly this corresponds to hierarchical states, where the asymptotic distribution among losing agents again exhibits a weak monopolist (see Section 3.3 for details).

2.2 Review of previous results

As already hinted in the introduction, the literature on generalized Pólya urns is vast. In this section, we shortly present a selection of results related to this work. The probably most widely known result concerns the limiting distribution of the classical Pólya urn, which traces back to [62, 23].

Theorem 2.4. *[24, Theorem 1] If $F_1(k) = \ldots = F_A(k) = k$, then $\lim_{n\to\infty}\chi(n)$ exists almost surely and has a Dirichlet distribution with parameter $X(0)$.*

[66] shows that the rate of convergence (in distribution) is $O(n^{-1})$ in Wasserstein distance. A powerful tool to study the linear Pólya urn is the theory of exchangeable processes, including de Finetti's famous theorem. Unfortunately, this is not applicable for our non-linear extension (see [74]). To our knowledge, the following Theorem 2.5 is the most comprehensive result concerning the long-time limit of our non-linear process $(\chi(n))_n$, which extends an earlier result in [73].

Theorem 2.5. *[30, Theorem 3.1] Suppose that $p(x) := \lim_{k\to\infty} p(k,x)$ (cf. (2.2)) exists for all $x \in \Delta_{A-1}$ and that even*

$$\sum_{k=1}^{\infty} \frac{\sup_{x\in\Delta_{A-1}} \|p(k,x) - p(x)\|}{k} < \infty \tag{2.7}$$

holds. Moreover, assume that there is a twice differentiable Lyapunov function for the vector field $(G(x))_{x\in\Delta_{A-1}} = (p(x) - x)_{x\in\Delta_{A-1}}$. Then $\chi(n)$ converges almost surely for $n \to \infty$ and the limit is in $\{x \in \Delta_{A-1} : G(x) = 0\}$.

Note that a Lyapunov function does always exist in the case $A = 2$ and when p is differentiable with equal feedback functions for all agents. Moreover, [30] shows under mild technical assumptions that each stable fixed point of the vector field G is attained in the limit $n \to \infty$ with positive probability, whereas unstable fixed points are never attained.

Theorem 2.5 allows to compute the long-time market shares in generic situations, like $F_i(k) = \alpha_i k^\beta$. Nevertheless, condition (2.7) is not fulfilled e.g. for $F_i(k) = \log(k)$, $F_i(k) = k\log(k)$ or $F_i(k) = e^k$.

In the monopoly case described in Theorem 2.2, the monopolist is in general random. Consequently, one is interested in the probability that a specific agent is the monopolist, at least in the limit $N \to \infty$. [98] derives such a result in a situation with only two symmetric agents.

Theorem 2.6. *[98, Theorem 2] Let $A = 2$ and $F_1 = F_2 = F$. Assume that F fulfills (M) and that*

$$\liminf_{x\to\infty} x\frac{d}{dx}\log F(x) > \frac{1}{2} \quad \text{and} \quad \lim_{x\to\infty} \frac{d}{dx}\log F(x) = 0.$$

Moreover, suppose that there is a constant $C > 0$ such that for all $\epsilon \in (0, \frac{1}{2})$ and all $x > 0$ large enough

$$\sup_{x\le t\le x^{1+\epsilon}} \left| \frac{t\frac{d}{dt}\log F(t)}{x\frac{d}{dx}\log F(x)} - 1 \right| \le C\epsilon$$

holds. Let $X(0) = (N + \lambda q(N), N - \lambda q(N))$ for $N, \lambda > 0$ and $q(a) := \sqrt{\frac{a}{4a\frac{d}{da}\log F(a) - 2}}$. Then the probability of agent 1 being the monopolist converges to $\Phi(\lambda)$ for $N \to \infty$, where Φ denotes the cumulative distribution function of the normal distribution.

For $F(x) = x^\beta$, $\beta > 1$ these assumptions are fulfilled and $q(N) = \frac{\sqrt{N}}{\sqrt{4\beta - 2}}$ is of order $\sqrt{N}$. This means that even a small initial advantage (compared to N) of one agent leads to this agent being the monopolist with high probability. Note that the assumptions of Theorem 2.6 are not satisfied for exponentially increasing feedback. Moreover, [102] shows under similar assumptions that $\mathbb{P}(sMon_i(\chi(0), N))$ is even exponentially decreasing in N when $\chi_i(0) < \frac{1}{2}$ for an agent i.

Theorem 2.7. *[102, Theorem 3] Let $A = 2$ and $F_1 = F_2 = F$. Assume that F fulfills (M) and that*

$$\liminf_{x\to\infty} x\frac{d}{dx}\log F(x) > 1 \quad \text{and} \quad \lim_{x\to\infty} x^{\frac{3}{4}}\frac{d}{dx}\log F(x) = 0.$$

Moreover, suppose that there is a constant $C > 0$ and $x_0 > 0$ such that for all $\epsilon \in (0,1)$ and all $x \geq x_0$

$$\sup_{x \leq t \leq x^{1+\epsilon}} \left| \frac{t \frac{d}{dt} \log F(t)}{x \frac{d}{dx} \log F(x)} - 1 \right| \leq C\epsilon$$

holds. If $0 < \chi_1(0) < x_1 < \frac{1}{2}$, then there exists $\gamma > 0$ such that

$$\mathbb{P}(\exists n \geq 0 \colon \chi_1(n) > x_1) \leq e^{-N^\gamma}.$$

[50] provides results for the asymmetric case $F_i(x) = \alpha_i x^\beta$, $i \in \{1,2\}, \beta > 1, \alpha_i > 0$. Under the assumptions of Theorem 2.7, [102] shows that the number of steps, in which the loser wins, has a heavy-tailed distribution. Here, for sequences $(x_n)_n$ and $(y_n)_n$ we write $x_n \sim y_n$ if $x_n/y_n \to 1$ for $n \to \infty$.

Theorem 2.8. *[102, Theorem 4] Under the assumptions of Theorem 2.7, the following holds:*

$$\mathbb{P}(\min\{X_1(\infty), X_2(\infty)\} > x) \sim \textit{const.} \sum_{k=x}^{\infty} \frac{1}{F(k)} \quad \textit{for } x \to \infty$$

[44, 132] show similar results on the number of steps won by the loser. Moreover, they computed the tail of the time when the monopoly occurs. We will generalize their achievements in Chapter 5.

Theorem 2.9. *[44, Theorem 10] Let $A = 2$ and $F_1(k) = F_2(k) = k^\beta$ with $\beta > 1$. Define the time of monopoly as*

$$N_{mon} \coloneqq \min\{n \geq 0 \colon \min\{X_1(n), X_2(n)\} = \min\{X_1(\infty), X_2(\infty)\}\}$$

and $\beta' \coloneqq \beta - \frac{\beta-1}{\beta}$. Then there are constants $c, C > 0$ such that

$$\frac{c}{n^{\beta'}} \leq \mathbb{P}(N_{mon} = n) \leq \frac{C}{n^{\beta'}}.$$

[44] also computes the tail of N_{mon} for other explicit feedback functions, but does not provide a handsome general expression. More recently in [97], a result for polynomial feedback with different exponents was shown.

Theorem 2.10. *[97, Theorem 2.2] Let $A = 2$ and $F_i(k) = k^{\beta_i}$ with $1 < \beta_1 \leq \beta_2$. Define the critical values*

$$\alpha_{cr} = \frac{\beta_1 - 1}{\beta_2 - 1} \quad \textit{and} \quad \nu_{cr} = \alpha_{cr}^{\frac{1}{\beta_2 - 1}}.$$

Moreover, set $X(0) = (x, \nu x^\alpha + o(x^\alpha))$ for $\alpha \in (0,1), \nu > 0$.

1. *If either $\alpha < \alpha_{cr}$ or $\alpha = \alpha_{cr}$ and $\nu < \nu_{cr}$, then $\lim_{x\to\infty} \mathbb{P}(sMon_1(X(0)) = 1$.*
2. *If either $\alpha > \alpha_{cr}$ or $\alpha = \alpha_{cr}$ and $\nu > \nu_{cr}$, then $\lim_{x\to\infty} \mathbb{P}(sMon_2(X(0)) = 1$.*

Hence, even the agent with inferior feedback can be the monopolist with high probability when there is a strong enough imbalance in the initial composition of the urn, i.e. $X_2(0) \sim X_1(0)^\alpha$ for $\alpha < \alpha_{cr}$. In addition, [97] provides a result for the critical case $\alpha = \alpha_{cr}$, $\nu = \nu_{cr}$.

For $F_i(k) = k^\beta$, $i \in [A]$ with $\beta < 1$, we know from Theorem 2.5 that $\lim_{n\to\infty} \chi_i(n) = \frac{1}{A}$ almost surely for all $i \in [A]$ irrespective of the initial configuration $\chi(0)$. The rate of convergence is specified in [84].

Theorem 2.11. *[84, Propsition 3] Let $F_i(k) = k^\beta$ for all $i \in [A]$ and $\beta \in (0, 1)$.*

1. *If $\frac{1}{2} < \beta < 1$, then*

$$n^{1-\beta}\left(\chi(n) - \frac{1}{A}\right) \xrightarrow{n\to\infty} C \quad \textit{almost surely}$$

   for a random, nonzero vector C.

2. *If $0 < \beta < \frac{1}{2}$ and $i \in [A]$, then*

$$\sqrt{n}\left(\chi_i(n) - \frac{1}{A}\right) \xrightarrow{n\to\infty} \mathcal{N}\left(0, \frac{A-1}{A^{1+2\beta}(1-2\beta)}\right) \quad \textit{in distribution,}$$

   where $\mathcal{N}$ denotes a Gaussian distribution.

3. *If $\beta = \frac{1}{2}$ and $i \in [A]$, then*

$$\sqrt{\frac{n}{\log(n)}}\left(\chi_i(n) - \frac{1}{A}\right) \xrightarrow{n\to\infty} \mathcal{N}\left(0, \frac{A-1}{A^2}\right) \quad \textit{in distribution.}$$

The convergence in part 2 and 3 can be extended to the vector $\chi(n)$. Part 1 implies that the leading agent does only change finitely often. According to [101, Theorem 1], this happens in general if and only if

$$\sum_{k=1}^{\infty} \frac{1}{F(k)^2} < \infty,$$

where $F = F_1 = \ldots = F_A$ fulfills $\liminf_{k\to\infty} F(k) > 0$. Also see [103] for details on that phenomenon. In the case $A = 2$ and $F_i(k) = \alpha_i k^\beta$, $\beta \geq 0$, [81] derives the tail distributions of the number and last times of ties $X_1(n) = X_2(n)$.

Recently, researchers were also interested in fluctuations of Pólya urns.

Theorem 2.12. *[39, Theorem 1.4] Let $A = 2$ and $F_1(k) = F_2(k) = \alpha k$ with $\alpha > 0$. Write $\chi^{(N)}(n) := \chi(n)$ to emphasize the dependence on the initial market size N and keep $\chi^{(N)}(0) = \chi(0)$ fixed for all N. Then the sequence of processes*

$$[0, \infty) \to [0, 1], \quad t \mapsto \sqrt{N}\left(\chi_1^{(N)}(\lfloor Nt \rfloor) - \chi_1(0)\right)$$

converges for $N \to \infty$ weakly (in the Skorokhod-space) to a time-changed Brownian motion.

[28] extends this result to general A. In Chapter 4, we derive a corresponding central limit theorem for non-linear Pólya urns. Finally, [6] adds a law of iterated logarithm and [61] derives an interesting sample-path large deviation principle.

2.3 Overview and main contributions of this work

One purpose of this work is to provide a comprehensive account of the theoretical properties of the non-linear Pólya urn model, which applies for a large class of feedback functions. As shown in Section 2.2, many results on this model have been derived before, most of which are restricted to the symmetric two-agent case or to polynomial feedback. While dropping these assumptions, we reveal some new facets of the process. The second purpose is to give a detailed example of how this model can be used to explain real world phenomena.

Chapter 3 deals with the long-time limiting behaviour of the process. In the case of strong monopoly (see Theorem 2.2), we prove that it is possible to predict the monopolist in most situations in the limit for large initial market size N. This dissects the space of initial shares Δ_{A-1} into attraction domains of the agents and our approach allows an explicit computation in generic examples. We also discuss the behaviour on the boundary between attraction domains as well as some differences between polynomially and exponentially increasing feedback. For sub-linear feedback, we provide a new formula for the asymptotics of the share process $(\chi(n))_n$, which even holds when the limit $\lim_{n\to\infty} \chi(n)$ does not exist. We particularly focus on the transition from sub-linear to super-linear feedback, where the system exhibits interesting behaviour including hierarchical states and random weak monopoly, which is not covered by previous results to our knowledge. In all cases, we discuss the sensitivity of the process to heterogeneity and to small changes of the feedback function, which is particularly strong for feedback close to the identity.

In Chapter 4, we derive a scaling limit for the full time evolution of market shares and characterize the fluctuations in a functional central limit theorem, where the limiting process is a Gaussian exchange process. This part uses techniques from stochastic approximation (see e.g. [107, 27, 100]) and extends previous results like Theorem 2.12 to non-linear feedback.

The goal of Chapter 5 is to extend Theorem 2.8 and related findings from [44, 132] on the wealth of losers to general, asymmetric feedback and to systems with more than two agents. For that, we first derive a result on the tail distribution of an explosive birth process at a random or fixed observation time. As a seeming paradox, losers with feedback close to the identity are most likely to win in many steps. Moreover, our approach allows insights in the dependencies of several losers, so that tails of other quantities can be derived, too. These are in particular the total wealth of all losers and the time of monopoly. In addition, we gain some new insights in the large deviation behaviour of non-linear Pólya urns. We also discuss the wealth of agents with sub-linear feedback as well as special features of large systems. Again, we particularly shed light on the behaviour close to the transition given by (M), i.e. between super- and sub-linear feedback.

Finally, in Chapter 6, we turn to rather applied questions. As a widely observed phenomenon, wealth is distributed significantly more unequal than wages. We propose a new extension of the non-linear Pólya urn model in order to explain the mechanism leading to this observation. In that model, agents accumulate wealth via savings on wages and via capital returns. After a mathematical analysis of this new model and fitting the parameters, we execute simulations, which accurately reproduce German empirical data from present and past. Assuming that this model is appropriate for describing the dynamics of wealth, we make some surprising predictions for the future, which turn out to be quite sensitive on presumed future interest rates. Our simulations implement the idea of increasing return rates, but we also explain within our model why different investment skills pose a less accurate explanation for the gap between wage and wealth distribution.

Chapter 3

Long-time Limit Theorems

It is well known (Theorem 2.2 or Theorem 2.5) that Pólya urns with non-linear feedback exhibit either strong monopoly or deterministic limit points. In the monopoly case, we present in Section 3.1 an asymptotic result for large market sizes on the distribution of the winner. For that, we have to distinguish exponential and polynomial feedback, which behave intrinsically differently, e.g. with regard to the sensitivity to unequal fitness of agents. In the non-monopoly case, we present in Section 3.2 a novel approach to compute the deterministic long-time market shares, which do not depend on the initial condition or early dynamics. In Section 3.3, we study in detail the transition between strong monopoly and deterministic limiting behaviour for almost linear feedback functions, which are particularly relevant in various applications including wealth dynamics (see Chapter 6). Predictable behaviour can only be expected for large initial market size, the behaviour of very young markets is intrinsically random. While bounds on the probabilities of certain events can be obtained, we focus here mostly on asymptotic results. To our knowledge, this provides the most complete account of the possible long-time behaviour for the generalized non-linear Pólya urn. In Section 3.4, we provide again a detailed summary of the various possible regimes of the process.

The results of this chapter have been published in [67].

3.1 Asymptotics for the strong monopoly case

Throughout this section, we assume that at least one agent i fulfills Condition (M), so that strong monopoly occurs with probability one. To characterize the asymptotics, we have to distinguish two different types of feedback functions with slightly different behaviour.

Definition 3.1. Let agent i (or F_i) fulfill (M). We call i (or F_i) of **type P** (for polynomial) if

$$\lim_{k\to\infty} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} = \infty \tag{P}$$

and of **type E** (for exponential) if

$$\limsup_{k\to\infty} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} < \infty. \tag{E}$$

For the rest of this section we assume that all agents with feedback functions that fulfill (M) are either of type P or type E. Of course it is possible to construct counter-examples (see Example 3.3), but these two types still cover a very large range, including most previous results.

Proposition 3.2. *If*

$$\frac{d}{dx}\log(F(x)) \xrightarrow{x\to\infty} 0 \tag{3.1}$$

then F is of type P, and if

$$\liminf_{x\to\infty} \frac{d}{dx}\log(F(x)) > 0 \tag{3.2}$$

then F is of type E.

Proof. First we assume (3.1) and observe that

$$\frac{F(k+1)}{F(k)} = \exp\left\{\int_k^{k+1} \frac{d}{dx}\log(F(x))dx\right\} \xrightarrow{k\to\infty} 1. \tag{3.3}$$

Consequently, for any given $\epsilon > 0$ there exists k_0 such that $\forall k \geq k_0 : F(k+1)/F(k) \leq 1+\epsilon$. Then we get for $k \geq k_0$:

$$F(k)\sum_{l=k}^{\infty}\frac{1}{F(l)} = \sum_{l=k}^{\infty}\prod_{m=k+1}^{l}\frac{F(m-1)}{F(m)} \geq \sum_{l=k}^{\infty}\left(\frac{1}{1+\epsilon}\right)^{l-k} = \frac{1}{1-\frac{1}{1+\epsilon}} \xrightarrow{\epsilon\to 0} \infty$$

The result for type E follows similarly. □

This means that functions that grow exponentially or faster are of type E, whereas functions that grow slower than exponential (like polynomials) are of type P. Note that Oliveira's "valid feedback functions" in [102] or [101] are of type P, which includes furthermore all regular varying functions.

Example 3.3.

1. The conditions from Proposition 3.2 are not necessary for being type P resp. E. For instance take any function F of type E and define $\tilde{F}(2k) = \tilde{F}(2k+1) = F(k)$. Then $\tilde{F}$ is also of type E, but does obviously not fulfill (3.2).

2. A possible construction of a feedback function that is neither of type P nor type E, but satisfies (M), is the following. Take a function F such that

$$0 < \lim_{k\to\infty} F(k)\sum_{l=k}^{\infty}\frac{1}{F(l)} < \infty.$$

   holds, e.g. $F(k) = e^k$. Then define a new feedback function $\tilde{F}$ by replacing each $F(k)$ by k elements that all equal $kF(k)$, i.e.

$$\left(\tilde{F}(1), \tilde{F}(2), \dots\right) = \Big(F(1), 2F(2), 2F(2), 3F(3), 3F(3), 3F(3), \dots\Big).$$

   One can easily check that $\tilde{F}_i$ satisfies neither (E) nor (P).

3.1.1 Asymptotic attraction domains

If at least one agents fulfills the monopoly condition (M), we know by Theorem 2.2 that there is a strong monopoly, where all agents satisfying (M) have a positive probability of being the monopolist. Thus, the monopolist is in general random. Nevertheless, in most situations it is possible to predict the winner with high probability for large initial market size.

Definition 3.4. The **asymptotic attraction domain** of an agent $i \in [A]$ is defined as

$$D_i = \left\{ \chi(0) \in \Delta^o_{A-1} \colon \lim_{N\to\infty} \mathbb{P}(sMon_i(\chi(0), N)) = 1 \right\} \subset \Delta^o_{A-1}.$$

Obviously, the asymptotic attraction domains are disjoint. The main result of this section states that the asymptotic attraction domains cover the whole simplex up to boundaries under mild regularity conditions.

Theorem 3.5. *Let at least one agent satisfy (M) and all agents satisfying (M) are either of type P or type E. Moreover, assume that one of the following conditions holds:*

1. *At least one agent is of type E and for all $\chi(0) \in \Delta^o_{A-1}$, $i, j \in [A]$*

$$\liminf_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = 0 \quad \text{and} \quad \limsup_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \infty \quad \text{do not hold simultaneously} .$$

2. *No agent is of type E and all agents of type P (there is at least one) fulfill*

$$\limsup_{k\to\infty} \frac{1}{k} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} < \infty. \tag{3.4}$$

   In addition, suppose that $\lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} \in [0,\infty]$ exists for all $\chi(0) \in \Delta^o_{A-1}, i, j \in [A]$.

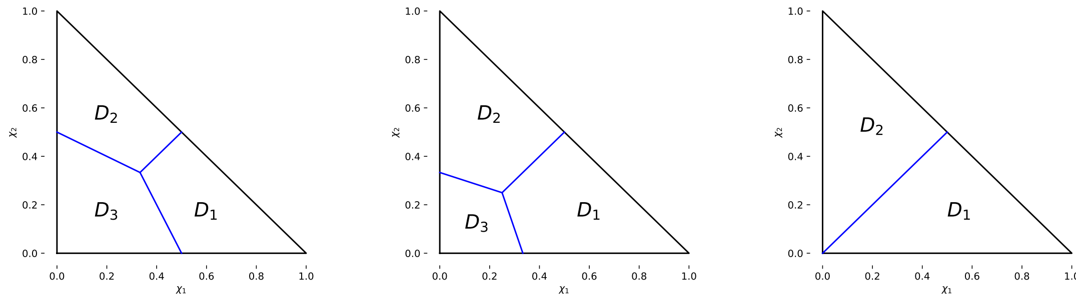
Then the asymptotic attraction domains are polytopes that dissect the simplex up to boundaries, i.e.

$$\bigcup_{i=1}^{A} \overline{D_i} = \Delta_{A-1},$$

where $\overline{(\cdot)}$ is the topological closure. If agent i does not satisfy (M)*, then $D_i = \emptyset$.*

As a direct consequence of the exponential embedding, $\mathbb{P}(sMon_i(\chi(0), N)) = 0$ for all agents that do not fulfill (M). Hence, their attraction domains are empty. Moreover, the attraction domains of agents of type P are empty if there is an agent of type E (Proposition 3.10). Theorem 3.5 basically follows from the results presented in the following subsections, where e.g. explicit conditions for

$$\lim_{N\to\infty} \mathbb{P}(sMon_i(\chi(0), N)) = 1$$



(a) $F_1(k) = F_2(k) = F_3(k) = k^2$ (b) $F_1(k) = F_2(k) = 2k^2$, $F_3(k) = k^2$ (c) $F_1(k) = F_2(k) = k^3$, $F_3(k) = k^2$

Figure 3.1: Asymptotic attraction domains in the case $A = 3$ with various feedback functions.

as well as bounds of $\mathbb{P}(sMon_i(\chi(0), N))$ are derived. The final proof of Theorem 3.5 will be presented in Subsection 3.1.4. It will turn out that the explosion times T_i from Section 2.1 concentrate on their expectations, i.e.

$$\lim_{N\to\infty} \frac{T_i(N\chi_i(0))}{\mathbb{E}T_i(N\chi_i(0))} = 1 \quad \text{almost surely,}$$

only for agents of type P, but not for type E, so we need to study these two types of feedback functions separately. The technical conditions in each case are mild and will be discussed in the following subsections. For example, Condition (3.4) excludes feedback close to the identity like $F_i(k) = k(\log k)^\beta$, $\beta > 1$, but this condition is in general not necessary (see Example 3.19). Another characteristic of type E is, that a strong monopoly is typically even a total monopoly, at least when N is large.

Theorem 3.6. *Let all agents be of type E and suppose that Assumption 1 in Theorem 3.5 is satisfied. If $\chi(0) \in D_i^o$ is in the interior of D_i for some $i \in [A]$, then*

$$\lim_{N\to\infty} \mathbb{P}(tMon_i(\chi(0), N)) = 1. \tag{3.5}$$

Theorem 3.6 is a direct consequence of Theorem 3.7 given below. As explained in Corollary 3.12, total monopoly does in general not occur, if $\chi(0)$ is on the boundary of the attraction domain. In addition, it turns out that in generic situations the probability of total monopoly is bounded away from one if all agents are of type P (Example 3.9). As explained in Example 3.20, the boundary between attraction domains can also belong to one of them.

3.1.2 Agents of type E and total monopoly

This subsection examines the occurrence of strong monopoly for diverging initial market size. It turns out that total monopoly of an agent $i \in [A]$ occurs with high probability (for $N \to \infty$) if either i is of type E and (3.6) holds, or if (3.8) holds. If at least one agent is of type E, then the attraction domains of all agents of type P are empty. On the edge between attraction domains, total monopoly occurs with probability one only for super-exponentially growing feedback.

The following results basically imply the first part of Theorem 3.5 as well as Theorem 3.6 as described in Subsection 3.1.4. The main result of this subsection provides a useful lower and upper bound for the probability of total monopoly.

Theorem 3.7. *Let agent $i \in [A]$ fulfill (M). Then for all $\chi(0) \in \Delta^o_{A-1}$ and $N \geq A$*

$$\prod_{j\neq i} \exp\left\{-F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}\right\} \leq \mathbb{P}(tMon_i(\chi(0), N)) \leq \prod_{j\neq i} \exp\left\{-c_N F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}\right\}$$

where

$$c_N := \inf_{k\in\mathbb{N}_0} \frac{F_i(\chi_i(0)N+k)}{F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)} > 0 \ .$$

Proof. Direct calculation yields

$$\mathbb{P}(tMon_i(\chi(0), N)) = \prod_{k=0}^{\infty} \frac{F_i(\chi_i(0)N+k)}{F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)}$$
$$\geq \exp\left\{-\sum_{k=0}^{\infty} \frac{\sum_{j\neq i} F_j(\chi_j(0)N)}{F_i(\chi_i(0)N+k)}\right\} = \prod_{j\neq i} \exp\left\{-F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}\right\}$$

using the inequality $e^{-x} \leq \frac{1}{1+x}$ for $x > -1$. For the upper bound, we estimate

$$\mathbb{P}(tMon_i(\chi(0), N)) = \prod_{k=0}^{\infty} \frac{F_i(\chi_i(0)N+k)}{F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)}$$
$$= \exp\left\{\sum_{k=0}^{\infty}\left[\log(F_i(\chi_i(0)N+k) - \log\left(F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)\right)\right]\right\}$$
$$\overset{\star\star}{\leq} \exp\left\{-\sum_{k=0}^{\infty} \frac{\sum_{j\neq i} F_j(\chi_j(0)N)}{F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)}\right\}$$
$$= \prod_{j\neq i} \exp\left\{-F_j(\chi_j(0)N) \sum_{k=0}^{\infty} \frac{1}{F_i(\chi_i(0)N+k) + \sum_{j\neq i} F_j(\chi_j(0)N)}\right\}$$
$$\leq \prod_{j\neq i} \exp\left\{-c_N F_j(\chi_j(0)N) \sum_{k=0}^{\infty} \frac{1}{F_i(\chi_i(0)N+k)}\right\}$$

using $\log(x+y) - \log(x) \geq \frac{y}{x+y}$ in $\star\star$. □

An immediate consequence of Theorem 3.7 is that for any agent fulfilling (M) the probability of a total monopoly is positive, but less than one. In addition, Theorem 3.7 reveals a significant behavioural difference between agents of type E and type P: whereas total monopoly is very likely for type E agents when the initial market size N is large, it is rather untypical for type P, which is explained in the following corollary and example.

Corollary 3.8. *1. If agent $i \in [A]$ is of type E , then for all $\chi(0) \in \Delta^o_{A-1}$ the following are equivalent:*

$$\lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \infty \quad \textit{for all } j \neq i \tag{3.6}$$

$$\lim_{N\to\infty} \mathbb{P}(tMon_i(\chi(0),N)) = 1. \tag{3.7}$$

2. If agent $i \in [A]$ fulfills (M), then for all $\chi(0) \in \Delta^o_{A-1}$

$$F_j(\chi_i(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)} \xrightarrow{N\to\infty} 0 \quad \textit{for all } j \neq i \tag{3.8}$$

is sufficient for (3.7). If in addition $F_i(k)$ is monotone for large k, (3.8) is equivalent to (3.7).

Proof. 1. If i is of type E, then (3.6) implies

$$F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)} \leq F_j(\chi_j(0)N) \frac{const.}{F_i(\chi_i(0)N)} \xrightarrow{N\to\infty} 0$$

using (E), and (3.7) follows from the lower bound of Theorem 3.7. The necessity of (3.6) follows from

$$\mathbb{P}(tMon_i(\chi(0),N)) \leq \frac{F_i(\chi_i(0)N)}{\sum_{j=1}^{A} F_j(\chi_j(0)N)} = \Big(1 + \sum_{j\neq i} \frac{F_j(\chi_i(0)N)}{F_i(\chi_j(0)N)}\Big)^{-1} .$$

2. (3.8) implies that the lower bound of Theorem 3.7 converges to one so that (3.7) holds. Now we assume that (3.8) does not hold. If $\frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)}$ does not converge to infinity for some $j \neq i$, then with 1., (3.7) cannot hold. Thus, we can assume (3.6) for all $j \neq i$, which implies $c_N \xrightarrow{N\to\infty} 1$ for the upper bound in Theorem 3.7 due to asymptotic monotonicity of $F_i(\chi_i(0)N + k)$ as $N \to \infty$. The upper bound then implies that $\mathbb{P}(tMon_i(\chi(0),N))$ does not converge to one. □

Example 3.9. 1. In the polynomial case $F_i(k) = \alpha_i k^{\beta_i}$ with $\alpha_i > 0, i \in [A]$ and $1 < \beta_1 \leq \ldots \leq \beta_A$ Condition (3.8) is equivalent to $\beta_A > \beta_{A-1} + 1$ for all $\chi(0) \in \Delta^o_{A-1}$. If $\beta_A = \beta_{A-1} + 1$, then

$$\lim_{N\to\infty} \mathbb{P}(tMon_A(\chi(0),N)) = \prod_{\substack{j=1,\ldots,A-1:\\ \beta_A=\beta_j+1}} \exp\left\{-\frac{\alpha_j}{\alpha_A}\left(\frac{\chi_j(0)}{\chi_A(0)}\right)^{\beta_A-1}\right\} \in (0,1)$$

since $c_N \xrightarrow{N\to\infty} 1$ in Theorem 3.7 and $\lim_{N\to\infty} \mathbb{P}(tMon_j(\chi(0), N)) = 0$ for $j \neq A$. If $\beta_A < \beta_{A-1} + 1$, in particular if $\beta_1 = \ldots = \beta_A$, then $\lim_{N\to\infty} \mathbb{P}(tMon_i(\chi(0), N)) = 0$ for all agents. This is also reflected in the large deviation behaviour of the process (Corollary E.4).

2. When $F_i(k) = \alpha_i e^{\beta_i k}$ for $\alpha_i > 0, \beta_i > 0, i \in [A]$, then Condition (3.6) is equivalent to $\beta_i \chi_i(0) > \beta_j \chi_j(0)$.

Remarkably for type E agents, if $F_i(k) = \alpha_i F(k)$ for all i and a function F fulfilling (E), then for large N the almost surely deterministic monopolist does not depend on the fitness-parameters α_i, but is only determined by the initial condition due to the strong feedback effect of type E functions.

Moreover, Theorem 3.7 provides information about the rate of convergence in (3.7). If agent i is of type E, then Theorem 3.7 states together with $1 + x \leq e^x$ and $\prod_{l=1}^k (1 - x_l) \geq 1 - \sum_{l=1}^k x_i$, $x_1, \ldots x_k \geq 0$

$$\mathbb{P}(tMon_i(\chi(0), N)) \geq \prod_{j\neq i} \left(1 - C\frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)}\right) \geq 1 - C\sum_{j\neq i} \frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)},$$

where

$$C := \sup_{k\geq 1} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} < \infty$$

because of (E). Thus, the convergence can be considered as quite fast, e.g. exponentially fast for $F_i(k) = e^k$.

Indeed, Condition (3.6) is fulfilled for an i in most generic cases, when at least one agent is of type E. To be more precise: If the expression in (3.6) neither tends to infinity nor to zero, then an arbitrarily small change in the initial market shares provides (3.6).

Proposition 3.10. *Let agent i be of type E.*

1. *If $j \neq i$ is of type P, then $\lim_{N\to\infty} \mathbb{P}(sMon_j(\chi(0), N)) = 0$ and $D_j = \emptyset$.*

2. *If $j \neq i$ is of type E and*

$$\liminf_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} > 0, \tag{3.9}$$

then for any $\epsilon > 0$:

$$\lim_{N\to\infty} \frac{F_i((\chi_i(0) + \epsilon)N)}{F_j(\chi_j(0)N)} = \infty$$

Proof. 1. By (E) we have for agent i of type E that

$$\frac{\sum_{l=k+1}^{\infty} \frac{1}{F_i(l)}}{\sum_{l=k}^{\infty} \frac{1}{F_i(l)}} = 1 - \frac{1}{F_i(k)\sum_{l=k}^{\infty} \frac{1}{F_i(l)}} < 1 - c \tag{3.10}$$

for some $c \in (0,1)$ and k large enough, thus the sequence $\left(\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}\right)_N$ converges to zero faster than $(1-c)^{N/\chi_i(0)}$. For an agent $j \neq i$ of type P we have by (P) for any $d > 0$ and k large enough

$$\frac{\sum_{l=k+1}^{\infty} \frac{1}{F_j(l)}}{\sum_{l=k}^{\infty} \frac{1}{F_j(l)}} = 1 - \frac{1}{F_j(k) \sum_{l=k}^{\infty} \frac{1}{F_j(l)}} > 1 - d\,.$$

Thus, the sequence $\left(\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}\right)_N$ converges to zero slower than $(1-d)^{N/\chi_j(0)}$. Together this yields

$$\frac{\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}}{\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}} \xrightarrow{N\to\infty} \infty \tag{3.11}$$

exponentially fast as d is arbitrarily small. Next, we express the event of strong monopoly via the explosion times (2.6). Take $\epsilon \in (0,1)$.

$$\begin{aligned}\mathbb{P}(sMon_j(\chi(0),N)) &\leq \mathbb{P}(T_j(\chi_j(0)N) < T_i(\chi_i(0)N)) \\ &\leq \mathbb{P}\left(T_j(\chi_j(0)N) < \epsilon\mathbb{E}T_j(\chi_j(0)N)\right) + \mathbb{P}(T_i(\chi_i(0)N) > \epsilon\mathbb{E}T_j(\chi_j(0)N))\end{aligned}$$

The first summand converges to zero for $N \to \infty$ according to Lemma 3.15. For the second summand, we apply Markov's inequality and (3.11):

$$\mathbb{P}(T_i(\chi_i(0)N) > \epsilon\mathbb{E}T_j(\chi_j(0)N)) \leq \frac{\mathbb{E}T_i(\chi_i(0)N))}{\epsilon\mathbb{E}T_j(\chi_j(0)N)} \xrightarrow{N\to\infty} 0$$

Hence, $\lim_{N\to\infty} \mathbb{P}(sMon_j(\chi(0),N)) = 0$.

2. Now let agent $j \neq i$ be of type E and assume (3.9). Then with (E):

$$\liminf_{N\to\infty} \frac{\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}}{\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}} \geq \liminf_{N\to\infty} const.\frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} > 0$$

Iterated application of estimate (3.10) yields

$$\frac{\sum_{l=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(l)}}{\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}} < (1-c)^{\lfloor \epsilon N \rfloor} \xrightarrow{N\to\infty} 0 \quad \text{for some } c \in (0,1)\ ,$$

and as a consequence

$$\frac{\sum_{l=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(l)}}{\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}} = \frac{\sum_{l=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(l)}}{\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}} \cdot \frac{\sum_{l=\chi_i(0)N}^{\infty} \frac{1}{F_i(l)}}{\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}} \xrightarrow{N\to\infty} 0$$

Then the claim follows via (E) and

$$\frac{F_i((\chi_i(0)+\epsilon)N)}{F_j(\chi_j(0)N)} \geq \frac{\sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}}{\sum_{l=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(l)}} \cdot \frac{1}{F_j(\chi_j(0)N) \sum_{l=\chi_j(0)N}^{\infty} \frac{1}{F_j(l)}} \xrightarrow{N\to\infty} \infty\,.$$

□

Corollary 3.8 implies that for any agent $i \in [A]$ of type E

$$\{\chi(0) \in \Delta^o_{A-1} \colon (3.6) \text{ holds}\} \subseteq D_i.$$

Due to Proposition 3.10, these sets are even equal up to boundaries under Assumption 1 of Theorem 3.5. Moreover, the first part of Proposition 3.10 states that the attraction domains of all agents of type P are empty, if there is at least one agent of type E. Recall that for finite N the probability of monopoly is positive for all agents satisfying (M). Nevertheless, $sMon_j$ can in general not be replaced by $tMon_j$ in 1. of Proposition 3.10 as underlined by the following example.

Example 3.11. Let $A = 2$ and $F_1(k) = e^k$ and

$$F_2(k) = \begin{cases} k^2 & \text{for } k \text{ even} \\ e^k & \text{for } k \text{ odd.} \end{cases}$$

It is easy to check that F_1 is of type E and F_2 is of type P, but (3.6) does not hold for $\chi_2(0) \geq \chi_1(0)$. As a consequence, $\lim_{N\to\infty} \mathbb{P}(sMon_1(\chi(0), N)) = 1$, but the monopoly is not necessarily total even for large N.

Finally, one can ask what happens for large N and **critical market shares**, i.e. for $\chi(0)$ lying exactly on the edge between the asymptotic attraction domains. It stands to reason that in this situation the monopolist remains random even for large N. Nevertheless, the exact limiting behaviour depends on whether the feedback functions grow exponentially or even super-exponentially.

Corollary 3.12. *Consider $\chi(0) \in \Delta^o_{A-1}$, such that there is a subset $S \subset [A]$ of agents with*

$$\forall i \in S,\, j \in [A] \colon \limsup_{N\to\infty} \frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)} < \infty \tag{3.12}$$

and

$$\forall j \notin S\, \exists i \in [A] \colon \limsup_{N\to\infty} \frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)} = 0\,. \tag{3.13}$$

Then the following holds:

1. *For all agents $i \in S$ of type E we have $\liminf_{N\to\infty} \mathbb{P}(tMon_i(\chi(0), N)) > 0$.*
2. *If for all agents $i \in S$ we have super-exponentially growing feedback, i.e.*

$$\lim_{k\to\infty} \frac{F_i(k+1)}{F_i(k)} = \infty,$$

then $\lim_{N\to\infty} \mathbb{P}\left(\bigcup_{i\in S} tMon_i(\chi(0), N)\right) = 1$.

3. *If for all agents $i \in S$ we have at most exponentially growing feedback, i.e.*

$$\limsup_{k\to\infty} \frac{F_i(k+1)}{F_i(k)} < \infty,$$

and $\#S \geq 2$, then $\limsup_{N\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{A} tMon_i(\chi(0), N)\right) < 1$.

Proof. 1. This follows directly from Theorem 3.7 and (E):

$$\mathbb{P}(tMon_i(\chi(0)) \geq \prod_{j\neq i} \exp\left\{-F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}\right\} \geq \prod_{j\neq i} \exp\left\{-const.\frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)}\right\}$$

2. First, we write

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{j=1}^{A} tMon_j(\chi(0), N)\right) \\
&= \sum_{i=1}^{A} \mathbb{P}\left(\bigcup_{j=1}^{A} tMon_j(\chi(0), N) \,\middle|\, X(1) - X(0) = e^{(i)}\right) \mathbb{P}\left(X(1) - X(0) = e^{(i)}\right) \\
&= \sum_{i=1}^{A} \mathbb{P}\left(tMon_i\left(\frac{1}{N+1}\left(\chi(0)N + e^{(i)}\right), N+1\right)\right) \mathbb{P}\left(X(1) - X(0) = e^{(i)}\right) \\
&\sim \sum_{i\in S} \mathbb{P}\left(tMon_i\left(\frac{1}{N+1}\left(\chi(0)N + e^{(i)}\right), N+1\right)\right) \mathbb{P}\left(X(1) - X(0) = e^{(i)}\right) \quad \text{for } N \to \infty\,,
\end{aligned}$$

where the last line is a consequence of (3.13). This also implies that in the limit $N \to \infty$ only agents in S can be the total monopolist. Note that super-exponential feedback implies being of type E (see Proposition 3.2). Then apply Theorem 3.7 and (E) for $i \in S$:

$$\begin{aligned}
&\mathbb{P}\left(tMon_i\left(\frac{1}{N+1}\left(\chi(0)N + e^{(i)}\right), N+1\right)\right) \geq \prod_{j\neq i} \exp\left\{-F_j(\chi_j(0)N) \sum_{k=\chi_i(0)N+1}^{\infty} \frac{1}{F_i(k)}\right\} \\
&\geq \prod_{j\neq i} \exp\left\{-const.\frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N+1)}\right\} = \prod_{j\neq i} \exp\left\{-const.\frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)} \cdot \frac{F_i(\chi_i(0)N)}{F_i(\chi_i(0)N+1)}\right\} \\
&\xrightarrow{N\to\infty} 1
\end{aligned}$$

3. Similarly to the second part, this follows from

$$\begin{aligned}
&\mathbb{P}\left(tMon_i\left(\frac{1}{N+1}\left(\chi(0)N + e^{(i)}\right), N+1\right)\right) \leq \frac{F_i(\chi(0)N+1)}{F_i(\chi_i(0)N+1) + \sum_{j\neq i} F_j(\chi_j(0)N)} \\
&= \left(1 + \sum_{j\neq i} \frac{F_j(\chi_j(0))}{F_i(\chi_i(0)N)} \cdot \frac{F_i(\chi_i(0)N)}{F_i(\chi_i(0)N+1)} \cdot\right)^{-1}.
\end{aligned}$$

The claim follows since $\#S \geq 2$ and $\liminf_{N\to\infty} \frac{F_j(\chi_j(0))}{F_i(\chi_i(0))} > 0$ for $j \in S$. □

Example 3.13. Let $F_i(k) = e^{\alpha_i k^\beta}$ for $\alpha_i > 0, \beta > 0$ and all $i \in [A]$. Then Condition (3.12) is satisfied with $S = [A]$ if $\alpha_i \chi_i(0)^\beta = \alpha_j \chi_j(0)^\beta$ for all $i, j \in [A]$. According to Corollary 3.12, we have in this case $\lim_{N\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{A} tMon_i(\chi(0), N)\right) = 1$ for $\beta > 1$ and $\limsup_{N\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{A} tMon_i(\chi(0), N)\right) < 1$ for $\beta \leq 1$.

3.1.3 Agents of type P

Let us now turn to the more widely studied case when all agents are of type P. We already saw in Example 3.9 that in this case a total monopoly is rather untypical. Since the definition of type P includes the monopoly condition (M), strong monopoly still occurs with probability one. This subsection prepares the proof of the second part of Theorem 3.5, which is presented in Subsection 3.1.4. In addition, we also discuss the possible behaviour on the boundary between attraction domains as well as feedback that is excluded from Theorem 3.5 via Condition (3.4).

We start with a result on the prediction of the monopolist in the limit $N \to \infty$.

Theorem 3.14. *Let all agents be of type P or not fulfill (M). If there is an agent $i \in [A]$ of type P such that*

$$\limsup_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} < 1 \quad \textit{for all } j \neq i \ , \tag{3.14}$$

then

$$\lim_{N\to\infty} \mathbb{P}(sMon_i(\chi(0), N)) = 1 \ . \tag{3.15}$$

Note that Condition (3.14) can be replaced by the easier, but stricter condition

$$\limsup_{N\to\infty} \frac{F_j(\chi_j(0)N)}{F_i(\chi_i(0)N)} < \frac{\chi_j(0)}{\chi_i(0)}$$

since

$$\limsup_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = \limsup_{N\to\infty} \frac{\int_{\chi_i(0)N}^{\infty} \frac{1}{F_i(u)} du}{\int_{\chi_j(0)N}^{\infty} \frac{1}{F_j(u)} du} = \limsup_{N\to\infty} \frac{\chi_i(0) \frac{1}{F_i(\chi_i(0)N)}}{\chi_j(0) \frac{1}{F_j(\chi_j(0)N)}}$$

due to de l'Hôspital's Theorem. This implies that for regular varying $F_i(k) = \alpha_i k^\beta L(k)$, where $\beta > 1$ and L is a slowly varying function, the attraction domains are equal to the polynomial case, where $F_i(k) = \alpha_i k^\beta$. Moreover, the attraction domains do not change if F_i is replaced by another function $\tilde{F}_i$ satisfying $\lim_{k\to\infty} \frac{\tilde{F}_i(k)}{F_i(k)} = 1$.

Proof. This is an immediate consequence of the following Lemma 3.15 and the exponential embedding representation (2.6) of the strong monopoly via

$$\mathbb{P}\left(T_i(\chi_i(0)N) < T_j(\chi_j(0)N)\right) = \mathbb{P}\left(\frac{T_i(\chi_i(0)N)}{\mathbb{E}T_i(\chi_i(0)N)} \cdot \frac{\mathbb{E}T_j(\chi_j(0)N)}{T_j(\chi_j(0)N)} \cdot \frac{\mathbb{E}T_i(\chi_i(0)N)}{\mathbb{E}T_j(\chi_j(0)N)} < 1 \right)$$
$$\xrightarrow{N\to\infty} 1,$$

since $\mathbb{E}T_i(\chi_i(0)N) = \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}$. □

Lemma 3.15. *If agent i is of type P, then:*

$$Var\left(\frac{T_i(\chi_i(0)N)}{\mathbb{E}T_i(\chi_i(0)N)} \right) \xrightarrow{N\to\infty} 0$$

Proof. We can find an appropriate regular extension of F_i, such that

$$\sum_{k=n}^{\infty} \frac{1}{F_i(k)} \sim \int_n^{\infty} \frac{dx}{F_i(x)} \quad \text{and} \quad \sum_{k=n}^{\infty} \frac{1}{F_i(k)^2} \sim \int_n^{\infty} \frac{dx}{F_i(x)^2} \quad \text{for } n \to \infty .$$

By the Theorem of de L'Hôspital and (P) this implies

$$\lim_{N\to\infty} Var\left(\frac{T_i(\chi_i(0)N)}{\mathbb{E}T_i(\chi_i(0)N)}\right) = \lim_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)^2}}{\left(\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}\right)^2} = \lim_{N\to\infty} \frac{\int_{\chi_i(0)N}^{\infty} \frac{dx}{F_i(x)^2}}{\left(\int_{\chi_i(0)N}^{\infty} \frac{dx}{F_i(x)}\right)^2}$$
$$= \lim_{N\to\infty} \frac{1}{2F_i(\chi_i(0)N) \sum_{k=\chi_i(0)N} \frac{1}{F_i(k)}} = 0 .$$

□

Example 3.16. If $F_i(k) = \alpha_i k^\beta$ for all $i \in [A]$ and $\beta > 1$, then Condition (3.14) is equivalent to $\alpha_i \chi_i(0)^{\beta-1} > \alpha_j \chi_j(0)^{\beta-1}$. Thus, in contrast to the type E case (Example 3.9), the attractiveness-parameters α_i affect the attraction domains.

Lemma 3.15 uncovers another behavioural difference between type P and type E agents: For type P agents the explosion time concentrates on its expectation, whereas the variance of $T_i(\chi_i(0)N)/\mathbb{E}T_i(\chi_i(0)N)$ remains bounded from below for type E agents by an analogous argument, using (E). For many type P agents, including $F_i(k) = \alpha_i k^{\beta_i}$, it is possible to prove that the convergence of $T_i(\chi_i(0)N)/\mathbb{E}T_i(\chi_i(0)N)$ is even almost sure (see the proof of Proposition 3.21 together with the Lemma of Borel-Cantelli).

It is now natural to look for an analogy to Proposition 3.10 for type P agents in order to make sure that (3.14) is fulfilled for almost all initial market shares $\chi(0)$. Unfortunately, this attempt is meant to fail as the example $F_i(k) = F_j(k) = k(\log k)^\alpha$ for $\alpha > 1$ shows. In this case

$$\sum_{k=\lfloor \chi_i(0)N \rfloor}^{\infty} \frac{1}{F_i(k)} \sim \int_{\chi_i(0)N}^{\infty} \frac{1}{x(\log x)^\alpha} dx = \frac{1}{\alpha - 1} \log(\chi_i(0)N)^{1-\alpha} .$$

Therefore, (3.14) is not fulfilled for all choices of $\chi_i(0), \chi_j(0)$, since

$$\lim_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = \lim_{N\to\infty} \left(\frac{\log(N) + \log(\chi_i(0))}{\log(N) + \log(\chi_j(0))}\right)^{1-\alpha} = 1.$$

Nevertheless, with a further condition we can find a similar result as Proposition 3.10.

Proposition 3.17. *Suppose that for some $i \neq j$*

$$\lim_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = 1. \tag{3.16}$$

1. *If there exists $C < \infty$ such that for all $k \in \mathbb{N}$*

$$\frac{1}{k} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} \leq C,$$

i.e. (3.4) holds, then for all $\epsilon > 0$

$$\limsup_{N\to\infty} \frac{\sum_{k=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} < 1 \ .$$

2. *If*

$$\lim_{k\to\infty} \frac{1}{k} F_i(k) \sum_{l=k}^{\infty} \frac{1}{F_i(l)} = \infty \quad \text{(i.e. } i \text{ is in particular of type } P\text{)} \tag{3.17}$$

and (3.16) holds for one choice of $\chi_i(0), \chi_j(0)$, then (3.16) holds for all choices $\chi_i(0), \chi_j(0) \geq 0$ with $\chi_i(0) + \chi_j(0) \leq 1$.

Proof. 1. We have by (3.4)

$$\frac{\sum_{k=\chi_i(0)N+1}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}} = 1 - \frac{1}{F_i(\chi_i(0)N) \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}} \leq 1 - \frac{1}{C\chi_i(0)N} \tag{3.18}$$

and iterated application of this yields

$$\frac{\sum_{k=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}} \leq \left(1 - \frac{1}{C(\chi_i(0)+\epsilon)N}\right)^{\lfloor \epsilon N \rfloor} \xrightarrow{N\to\infty} e^{-\frac{\epsilon}{C(\chi_i(0)N+\epsilon)}} < 1 \ .$$

Finally, this implies

$$\limsup_{N\to\infty} \frac{\sum_{k=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = \limsup_{N\to\infty} \frac{\sum_{k=(\chi_i(0)+\epsilon)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}} \cdot \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} < 1 \ .$$

2. The second part follows by similar arguments, using Condition (3.17) for an "$\geq$"-estimate in (3.18), where $C = C(N)$ is arbitrarily large. □

Example 3.18. For all $i \in [A]$, let $F_i(k) = \alpha_i F(k)$ with a feedback function F fulfilling (3.17), e.g. $F(k) = k\log(k)^\beta$ for $\beta > 1$. Note that this situation is not covered by Theorem 3.5. According to 2. of Proposition 3.17, we then have

$$\lim_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = \lim_{N\to\infty} \frac{\alpha_j \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F(k)}}{\alpha_i \sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F(k)}} = \frac{\alpha_j}{\alpha_i} \quad \text{for all } \chi(0) \in \Delta_{A-1}$$

and, hence, $D_i = \Delta_{A-1}^o$ if $\alpha_i > \alpha_j$ for all $j \neq i$. Hence for large market size, the agent with the highest fitness α_i will be the monopolist independently of the initial share.

If all agents are of type P, Theorem 3.14 implies that for any agent $i \in [A]$

$$\{\chi(0) \in \Delta^o_{A-1} \colon (3.14) \text{ holds}\} \subseteq D_i .$$

Assuming 2. in Theorem 3.5, we get from Proposition 3.17 that the sets are equal up to boundaries.

In the situation of the second part of Proposition 3.17, the explosion times concentrate asymptotically on the same value, i.e. $T_i(\chi_i(0)N)/T_j(\chi_j(0)N) \xrightarrow{N\to\infty} 1$ in distribution. Thus, it is not possible to predict the monopolist for large N by means of the general results of this section, but we can still cover this boundary case based on an example.

Example 3.19. Let $A = 2$ and $F(k) = F_1(k) = F_2(k) = k(\log k)^\beta$ for $\beta > 1$. Assume $\chi_1(0) > \chi_2(0)$. Then

$$\begin{aligned}\mathbb{E}T_2(N\chi_2(0)) - \mathbb{E}T_1(N\chi_1(0)) &= \sum_{k=\chi_2(0)N}^{\chi_1(0)N-1} \frac{1}{F(k)} \sim \frac{1}{\beta-1}\Big(\log(\chi_2(0)N)^{1-\beta} - \log(\chi_1(0)N)^{1-\beta}\Big) \\ &\sim \mathit{const.}(\log N)^{-\beta} \quad \text{for } N \to \infty ,\end{aligned}$$

but for the variance we have

$$Var(T_i(N\chi_i(0))) = \sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F(k)^2} \sim \frac{\mathit{const.}}{N(\log N)^{2\beta}} \quad \text{for } N \to \infty$$

due to Karamata's theorem. Using Chebyshev's inequality as in Example 3.20, we get that

$$\mathbb{P}(sMon_1(\chi(0), N)) = \mathbb{P}(T_2(N\chi_2(0)) > T_1(N\chi_1(0))) \xrightarrow{N\to\infty} 1 .$$

Hence, $\chi(0) \in D_1$ is in the attraction domain of agent 1. An analogous argumentation is possible for more general $F(k) = kL(k)$ with slowly varying L.

Let us now turn to the **behaviour on the edge between attraction domains**. If $F_1 = \ldots = F_A$ and $\chi_i(0) = \frac{1}{A}$ for all i, then $\mathbb{P}(sMon_i(\chi(0), N)) = \frac{1}{A}$ holds for all N for symmetry reasons, i.e. $\chi(0)$ does not belong to any attraction domain as the monopolist remains random even in the limit $N \to \infty$. The following example underlines that in some cases the boundary between attraction domains belongs to one of them.

Example 3.20. Consider the process for $A = 2$ and $F_1(k) = k^\beta$, $F_2(k) = \frac{k^\beta}{1+k^{-\delta}}$ for $\beta > 1$ and $\delta \in (0, \frac{1}{2})$. Then we have

$$\sum_{k=\chi_2(0)N}^{\infty} \frac{1}{F_2(k)} = \sum_{k=\chi_2(0)N}^{\infty} \left(\frac{1}{k^\beta} + \frac{1}{k^{\beta+\delta}}\right) \sim \sum_{k=\chi_2(0)N}^{\infty} \frac{1}{k^\beta} = \sum_{k=\chi_2(0)N}^{\infty} \frac{1}{F_1(k)}$$

and similarly

$$\sum_{k=\chi_2(0)N}^{\infty} \frac{1}{F_2(k)^2} \sim \sum_{k=\chi_2(0)N}^{\infty} \frac{1}{F_1(k)^2} \quad \text{for } N \to \infty .$$

Moreover, set $\chi_1(0) = \chi_2(0) = \frac{1}{2}$, such that (3.16) holds, and define $\epsilon_N := cN^{1-\beta-\delta}$ for a constant $c > 0$ as specified below. Chebyshev's inequality yields

$$\mathbb{P}\left(T_1(\chi_1(0)N) > \mathbb{E}T_1(\chi_1(0)N) + \epsilon_N\right) \leq \frac{Var(T_1(\chi_1(0)N))}{\epsilon_N^2} = \frac{\sum_{k=N/2}^{\infty} \frac{1}{F_1(k)^2}}{\epsilon_N^2} \xrightarrow{N\to\infty} 0$$

since $\sum_{k=N/2}^{\infty} \frac{1}{F_1(k)^2} \sim \frac{1}{2\beta-1}(N/2)^{1-2\beta}$ and $\delta < \frac{1}{2}$. Similarly,

$$\mathbb{P}\left(T_2(\chi_1(0)N) < \mathbb{E}T_2(\chi_1(0)N) - \epsilon_N\right) \leq \frac{Var(T_2(\chi_2(0)N))}{\epsilon_N^2} = \frac{\sum_{k=N/2}^{\infty} \frac{1}{F_2(k)^2}}{\epsilon_N^2} \xrightarrow{N\to\infty} 0.$$

In addition, we have for sufficiently large N and small $c > 0$ that

$$\mathbb{E}T_1(\chi_1(0)N) + \epsilon_N < \mathbb{E}T_2(\chi_2(0)N) - \epsilon_N,$$

because

$$\mathbb{E}T_2(\chi_2(0)N) - \mathbb{E}T_1(\chi_1(0)N) = \sum_{k=N/2}^{\infty} \frac{1}{k^{\beta+\delta}} \sim \frac{1}{\beta+\delta-1}(N/2)^{1-\beta-\delta}.$$

Thus for large N

$$\begin{aligned}
\mathbb{P}(sMon_1(\chi_1(0), N)) &= \mathbb{P}\big(T_1(\chi_1(0)N) < T_2(\chi_2(0)N)\big) \\
&\geq \mathbb{P}\Big(T_1(\chi_1(0)N) < \mathbb{E}T_1(\chi_1(0)N) + \epsilon_N \,\wedge\, T_2(\chi_2(0)N) > \mathbb{E}T_2(\chi_2(0)N) - \epsilon_N\Big) \\
&= \mathbb{P}\Big(T_1(\chi_1(0)N) < \mathbb{E}T_1(\chi_1(0)N) + \epsilon_N\Big)\mathbb{P}\Big(T_2(\chi_2(0)N) > \mathbb{E}T_2(\chi_2(0)N) - \epsilon_N\Big) \\
&\xrightarrow{N\to\infty} 1
\end{aligned}$$

using the independence of $T_1(\chi_1(0)N)$ and $T_2(\chi_1(0)N)$. Hence, $\chi(0) \in D_1$.

We finish this subsection with a result on the **rate of convergence** in (3.15). [50] presents a bound for $\mathbb{P}(sMon_i(\chi(0), N))$ in the case $F_i(k) = F_j(k) = k^\alpha$, but a generalization of this procedure is possible.

Proposition 3.21. *Let all agents be of type P with monotone feedback functions. If (3.14) and (3.1) hold for agent $i \in [A]$, i.e. $\chi(0) \in D_i$, we have*

$$\mathbb{P}(sMon_i(\chi(0), N)) \geq 1 - \sum_{j=1}^{A} \exp\left(-(d_j - \epsilon)\sqrt{F_j(\chi_j(0)N) \sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}}\right)$$

for any $\epsilon > 0$ and large enough N, where

$$d_j := g\left(\limsup_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}}\right) > 0 \quad \text{with} \quad g(x) := \frac{1-x}{1+x} \quad \text{for } j \neq i$$

and $d_i := \min_{j\neq i} d_j$.

This means that the rate of convergence in (P) gives a lower bound for the rate of convergence of $\mathbb{P}(sMon_i(\chi_i(0))$.

Proof. Once again, the proof uses the exponential embedding from Section 2.1. Let $t > 0$ and $s := \left(\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2}\right)^{-\frac{1}{2}}$. Then the Markov-inequality and monotone convergence yield for all $j \in [A]$ and $t > 0$:

$$\begin{aligned}
&\mathbb{P}\left(\sum_{l=k}^{\infty} \tau_j(l) - \sum_{l=k}^{\infty} \frac{1}{F_j(l)} > t\sqrt{\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2}}\right) \\
&\leq \exp\left(-s\sum_{l=k}^{\infty} \frac{1}{F_j(l)} - st\sqrt{\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2}}\right) \cdot \mathbb{E}\exp\left(s\sum_{l=k}^{\infty} \tau_j(l)\right) \\
&= \exp\left(-s\sum_{l=k}^{\infty} \frac{1}{F_j(l)} - t\right) \cdot \prod_{l=k}^{\infty} \mathbb{E}e^{s\tau_j(l)} = \exp\left(-s\sum_{l=k}^{\infty} \frac{1}{F_j(l)} - t\right) \cdot \prod_{l=k}^{\infty}\left(1 + \frac{s}{F_j(l) - s}\right) \\
&\leq \exp\left(-s\sum_{l=k}^{\infty} \frac{1}{F_j(l)} - t\right) \cdot \prod_{l=k}^{\infty} \exp\left(\frac{s}{F_j(l) - s}\right) = \exp\left(s\sum_{l=k}^{\infty}\left(-\frac{1}{F_j(l)} + \frac{1}{F_j(l) - s}\right) - t\right) \\
&\leq \exp\left(s^2\sum_{l=k}^{\infty} \frac{1}{(F_j(l) - s)^2} - t\right) \leq \exp\left(\frac{s^2}{\left(1 - \frac{s}{F_j(k)}\right)^2}\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2} - t\right) = c_j(k)e^{-t},
\end{aligned}$$

where $c_j(k) := \exp\left(\frac{1}{\left(1 - \frac{s}{F_j(k)}\right)^2}\right)$. Setting

$$t = (d_i - \epsilon)\frac{\sum_{l=k}^{\infty} \frac{1}{F_j(l)}}{\sqrt{\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2}}}$$

(which is positive for ϵ small enough since $d_j > 0$ for all $j \in [A]$) yields

$$\begin{aligned}
\mathbb{P}\left(\frac{\sum_{l=k}^{\infty} \tau_j(l)}{\sum_{l=k}^{\infty} \frac{1}{F_j(l)}} - 1 > (d_i - \epsilon)\right) &\leq c_j(k)\exp\left(-(d_i - \epsilon)\frac{\sum_{l=k}^{\infty} \frac{1}{F_j(l)}}{\sqrt{\sum_{l=k}^{\infty} \frac{1}{F_j(l)^2}}}\right) \\
&\leq c_j(k)\exp\left(-(d_i - \epsilon)\sqrt{F_j(k)\sum_{l=k}^{\infty} \frac{1}{F_j(l)}}\right).
\end{aligned}$$

The second estimate uses $F_j(l) \geq F_j(k)$ by monotonicity. Analogously, one can show

$$\mathbb{P}\left(\frac{\sum_{l=k}^{\infty} \tau_j(l)}{\sum_{l=k}^{\infty} \frac{1}{F_j(l)}} - 1 < -(d_j - \epsilon)\right) \leq e \cdot \exp\left(-(d_j - \epsilon)\sqrt{F_j(k)\sum_{l=k}^{\infty} \frac{1}{F_j(l)}}\right).$$

Both estimates then imply for large enough N together with (2.6):

$$\begin{aligned}
&\mathbb{P}(sMon_i(\chi(0),N)) \geq 1-\sum_{j\neq i}\Big(1-\mathbb{P}\big(T_i(\chi_i(0)N)<T_j(\chi_j(0)N)\big)\Big)\\
&\geq 2-A+\sum_{j\neq i}\mathbb{P}\left(\frac{T_i(\chi_i(0)N)}{\sum_{k=\chi_i(0)N}^{\infty}\frac{1}{F_i(k)}}<1+(d_j-\epsilon)\right)\cdot\mathbb{P}\left(\frac{T_j(\chi_j(0)N)}{\sum_{k=\chi_j(0)N}^{\infty}\frac{1}{F_j(k)}}>1-(d_j-\epsilon)\right)\\
&\geq 2-A+\left(1-c_i(\chi_i(0)N)\cdot\exp\left(-(d_i-\epsilon)\sqrt{F_i(\chi_i(0)N)\sum_{k=\chi_i(0)N}^{\infty}\frac{1}{F_i(k)}}\right)\right)\\
&\quad\cdot\sum_{j\neq i}\left(1-e\,\exp\left(-(d_j-\epsilon)\sqrt{F_j(\chi_j(0)N)\sum_{k=\chi_j(0)N}^{\infty}\frac{1}{F_j(k)}}\right)\right)\\
&\geq 1-(A-1)\,c_i(\chi_i(0)N)\cdot\exp\left(-(d_i-\epsilon)\sqrt{F_i(\chi_i(0)N)\sum_{k=\chi_i(0)N}^{\infty}\frac{1}{F_i(k)}}\right)\\
&\quad-\sum_{j\neq i}e\,\exp\left(-(d_j-\epsilon)\sqrt{F_j(\chi_j(0)N)\sum_{k=\chi_j(0)N}^{\infty}\frac{1}{F_j(k)}}\right)
\end{aligned}$$

In the last inequality we use $(1-x)(1-y)\geq 1-x-y$ for $x,y\geq 0$. Next, we observe that

$$c_i(k)=\exp\left(\frac{1}{\left(1-\frac{s}{F_i(k)}\right)^2}\right)=\exp\left(\frac{1}{\left(1-\frac{1}{F_i(k)\sqrt{\sum_{l=k}^{\infty}\frac{1}{F_i(l)^2}}}\right)^2}\right)\xrightarrow{k\to\infty} e\,,$$

because F^2 is also of type P due to (3.1). Using (P), this finally leads to

$$\mathbb{P}(sMon_i(\chi(0),N))\geq 1-\sum_{j=1}^{A}\exp\left(-(d_j-2\epsilon)\sqrt{F_j(\chi_j(0)N)\sum_{k=\chi_j(0)N}^{\infty}\frac{1}{F_j(k)}}\right)$$

for large enough N. $\square$

Using the notation $x_n \prec y_n$ if $\limsup_{n\to\infty}\frac{x_n}{y_n}<\infty$, we can rephrase Proposition 3.21 as

$$-\log\left(1-\mathbb{P}(sMon_i(\chi(0),N))\right)\succ\sum_{j=1}^{A}\sqrt{F_j(\chi_j(0)N)\sum_{k=\chi_j(0)N}^{\infty}\frac{1}{F_j(k)}}\quad\text{for } N\to\infty$$

if $\chi(0)\in D_i$ is in the attraction domain of agent i.

For $F_i(k)=\alpha k^{\beta}$, $\beta>1$, we have $F_i(k)\sum_{l=k}^{\infty}\frac{1}{F_i(l)}\sim\frac{k}{\beta-1}$, thus the convergence of $\mathbb{P}(sMon_i(\chi(0)N))\to 1$ can be considered as fast. Hence, $\mathbb{P}(sMon_i(\chi(0),N))$ is close to one even for moderate N, when $\chi(0) \in D_i$ is in the asymptotic attraction domain.

Further asymptotic results on strong monopoly, mainly in the type P case, can be found e.g. in [84, 98, 103, 50, 49, 97, 81].

3.1.4 Proof of Theorem 3.5 and Theorem 3.6

Finally, we shortly explain how Theorem 3.5 and Theorem 3.6 follow from the results of the previous subsections.

First, assume that Assumption 1 of Theorem 3.5 is satisfied, i.e. at least one agent is of type E. Then Corollary 3.8 implies that for any agent $i \in [A]$ of type E

$$\tilde{D}_i := \{\chi(0) \in \Delta^o_{A-1} \colon (3.6) \text{ holds}\} \subseteq D_i.$$

Obviously:

$$\tilde{D}_i = \bigcap_{j \neq i} \left\{ \chi(0) \in \Delta^o_{A-1} \colon \lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \infty \right\}$$

Due to 2. of Proposition 3.10, there is a ratio $r_{i,j} \in [0, \infty]$ such that

$$\lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \begin{cases} \infty & \text{if } \frac{\chi_i(0)}{\chi_j(0)} > r_{i,j} \\ 0 & \text{if } \frac{\chi_i(0)}{\chi_j(0)} < r_{i,j} \end{cases}$$

for each pair $i \neq j$ of agents. Note that

$$\lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \lim_{N\to\infty} \frac{F_i\left(\frac{\chi_i(0)}{\chi_j(0)}N\right)}{F_j(N)}.$$

Hence,

$$\tilde{D}_i = \bigcap_{j \neq i} \left\{ \chi(0) \in \Delta^o_{A-1} \colon \frac{\chi_i(0)}{\chi_j(0)} > r_{i,j} \right\}$$

is an intersection of half-spaces and the simplex, i.e. a polytope. Moreover, $\tilde{D}_1, \ldots, \tilde{D}_A$ cover the whole simplex up to boundaries, since the "winning"-relation $\lim_{N\to\infty} \frac{F_i(\chi_i(0)N)}{F_j(\chi_j(0)N)} = \infty$ is transitive. Thus, $D_1, \ldots, D_A$ cover the simplex up to boundaries as well and $\tilde{D}_i$ equals D_i up to boundaries. According to Corollary 3.8, we even have $\mathbb{P}(tMon_i(\chi(0), N)) \to 1$ for $N \to \infty$, if $\chi(0) \in \tilde{D}_i$ and all agents are of type E. Hence, Theorem 3.6 is proven, too.

If Assumption 2 of Theorem 3.5 is satisfied, the proof is analogous using Theorem 3.14 and Proposition 3.17. Note that

$$\lim_{N\to\infty} \frac{F_i(\chi_j(0)N)}{F_j(\chi_i(0)N)} = c \in [0, \infty] \quad \Longrightarrow \quad \lim_{N\to\infty} \frac{\sum_{k=\chi_i(0)N}^{\infty} \frac{1}{F_i(k)}}{\sum_{k=\chi_j(0)N}^{\infty} \frac{1}{F_j(k)}} = c\frac{\chi_i(0)}{\chi_j(0)}$$

due to the Theorem of de l'Hôspital.

In summary, for finite N the monopolist is random and even disadvantageous agents can win. If the initial market size N is large, it is possible to predict the winner with high probability depending on the initial market shares.

3.2 The non-monopoly case

Now we consider the case when no agent fulfills (M), such that no strong monopoly occurs. It is known that in the case of a standard Pólya urn, i.e. $F_i(k) = k$ for all agents, the limit $\chi(\infty) = \lim_{n\to\infty} \chi(n)$ exists almost surely and $\chi(\infty)$ has a Dirichlet-distribution with parameter $X(0)$ (Theorem 2.4). Thus, in the long run all agents have a stable, non-zero, random market share.

It is basically known (Theorem 2.5) that if the feedback functions grow significantly slower than linear, then $\chi(\infty)$ is deterministic. We present an alternative approach to the sub-linear case, which allows some additional insights and a simple calculation of the limiting point in generic examples. For example, particularly weak like $F_i(k) = \log(k)$ is not included in the results of [30] (Theorem 2.5). In addition, our approach allows to construct feedback functions such that $\chi(n)$ does not even converge for $n \to \infty$. In order to get deterministic limits, we will need a condition, which ensures that the feedback functions grow slow enough. We will mainly use:

$$\limsup_{k\to\infty} \frac{1}{k} F_i(k) \sum_{l=1}^{k} \frac{1}{F_i(l)} < \infty \tag{C1}$$

Note that this already implies that i does not fulfill (M). Feedback excluded by condition C1 will be discussed in Section 3.3 and Proposition 3.30. We add some examples to gain an understanding of this restriction.

Example 3.22.

1. (C1) is not satisfied if $\lim_{k\to\infty} \frac{F_i(k)}{k} = \infty$.
2. For $F_i(k) = k(\log(k+1))^\beta$ with $\beta \leq 1$ (C1) is not fulfilled as $\sum_{l=1}^{k} \frac{1}{F_i(l)} \sim \frac{(\log k)^{1-\beta}}{1-\beta}$ for $\beta < 1$.
3. If $F_i(k) = \alpha k^\beta$ for $\alpha > 0, \beta < 1$, then (C1) is fulfilled as $\sum_{l=1}^{k} \frac{1}{F_i(l)} \sim \frac{k^{1-\beta}}{\alpha(1-\beta)}$.
4. For $F_i(k) = \log(k+1)$ (C1) is fulfilled as $\sum_{l=1}^{k} \frac{1}{F_i(l)} \sim \frac{k}{\log(k)}$.
5. (C1) is fulfilled if $\liminf_{k\to\infty} F_i(k) > 0$ and $\limsup_{k\to\infty} F_i(k) < \infty$.

In fact, Condition (C1) contains a monotonicity in the following sense.

Proposition 3.23. *If F_i fulfills (C1) and for some $j \neq i$*

$$\limsup_{u\to\infty} \int_1^u \left(\frac{d}{dx} \log(F_j(x)) - \frac{d}{dx} \log(F_i(x)) \right) dx < \infty,$$

then F_j fulfills (C1), too.

Proof. The assumption implies via (3.3)

$$\frac{F_i(k+l)}{F_i(k)} \geq const. \frac{F_j(k+l)}{F_j(k)} \quad \text{for all } k, l \in \mathbb{N}.$$

Hence:

$$\frac{1}{k}F_i(k)\sum_{l=1}^{k}\frac{1}{F_i(l)} = \frac{1}{k}\sum_{l=1}^{k}\frac{F_i(k)}{F_i(l)} \geq const.\frac{1}{k}\sum_{l=1}^{k}\frac{F_j(k)}{F_j(l)} = const.\frac{1}{k}F_j(k)\sum_{l=1}^{n}\frac{1}{F_j(l)}$$

□

Proposition 3.23 implies in particular that if F_i satisfies (C1) and $\frac{d}{dx}\log(F_j(x)) \leq \frac{d}{dx}\log(F_i(x))$, then F_j satisfies (C1), too. This includes the case $F_j(k) = const.F_i(k)$.

In general, our approach even allows feedback functions that converge to zero as long as this convergence is not to fast, which is ensured by the following regularity condition:

$$\liminf_{k\to\infty}\frac{1}{k^p}F_i(k)\sum_{l=1}^{k}\frac{1}{F_i(l)} > 0 \quad \text{for some } p > \frac{1}{2} \tag{C2}$$

Note that (C2) is fulfilled for $F_i(k) = \alpha k^{\beta}$, $\beta \in \mathbb{R}$, but not for $F_i(k) = \alpha e^{-\beta k}$, $\beta > 0$. In analogy to Proposition 3.23 we get a monotonicity here in the sense that if F_i fulfills (C2) and

$$\limsup_{u\to\infty}\int_1^u\left(\frac{d}{dx}\log(F_i(x)) - \frac{d}{dx}\log(F_j(x))\right)dx < \infty,$$

then F_j fulfills (C2), too. We are now prepared for the main result of this section regarding the counting processes (2.3) of the exponential embedding from Section 2.1.

Theorem 3.24. *Let F_i fulfill (C1) and (C2). Then*

$$\frac{\Xi_i(t)}{a_i^{-1}(t)} \xrightarrow{t\to\infty} 1 \quad \textit{almost surely},$$

where a_i^{-1} denotes the inverse function of $a_i(t) := \int_1^t \frac{dx}{F_i(x)}$.

Note that a_i^{-1} exists as a_i is strictly monotone. We suppose that the extension $F_i\colon (0,\infty) \to (0,\infty)$ satisfies $a_i(k) \sim \sum_{l=1}^{k}\frac{1}{F_i(l)}$ for $k \to \infty$. The asymptotics of birth processes have been studied in the literature before (see Section 5.1 or Theorem 3.31). The following lemma provides the first step of the proof of Theorem 3.24, using standard ideas from renewal theory.

Lemma 3.25. *If F_i fulfills (C2) for $p > \frac{1}{2}$, then*

$$\frac{a_i(\Xi_i(t))}{t} \xrightarrow{t\to\infty} 1 \quad \textit{almost surely}.$$

Proof. (C2) implies

$$\lim_{k\to\infty}\sum_{l=1}^{k} Var\left(\frac{\tau_i(l)}{a_i(l)}\right) = \lim_{k\to\infty}\sum_{l=1}^{k}\frac{1}{F_i(l)^2 a_i(l)^2} \leq \lim_{k\to\infty}\sum_{l=1}^{k}\frac{const.}{l^{2p}} < \infty,$$

using $a_i(k) \sim \sum_{l=1}^{k} \frac{1}{F_i(l)}$ for $k \to \infty$. According to the Kolmogorov criterion (see e.g. [72, Section 6.2]) this is sufficient for

$$\frac{S_i(k)}{a_i(k)} \xrightarrow{k\to\infty} 1 \quad \text{almost surely,}$$

where $S_i(k) := \sum_{l=1}^{k} \tau_i(l)$. We use this and $\Xi_i(t) \to \infty$ $a.s.$ for the final estimate:

$$\begin{aligned} 1 = \lim_{t\to\infty} \frac{a_i(\Xi_i(t))}{S_i(\Xi_i(t)) - S_i(X_i(0)-1)} &\le \lim_{t\to\infty} \frac{a_i(\Xi_i(t))}{t} \le \lim_{t\to\infty} \frac{a_i(\Xi_i(t))}{S_i(\Xi_i(t)-1) - S_i(X_i(0)-1)} \\ &\le \lim_{t\to\infty} \frac{a_i(\Xi_i(t)-1)}{S_i(\Xi_i(t)-1)} = 1 \end{aligned}$$

In the second line, we used that $a_i(k-1) \sim a_i(k)$ obviously holds for $k \to \infty$. □

Now Theorem 3.24 is easy to prove.

Proof. Lemma 3.25 states

$$a_i(\Xi_i(t)) = t + o(t) \quad \Leftrightarrow \quad \Xi_i(t) = a_i^{-1}(t + o(t)) \quad \text{almost surely}$$

(using the Landau o-notation). It remains to show that

$$\lim_{t\to\infty} \frac{a_i^{-1}(t+o(t))}{a_i^{-1}(t)} = 1\,. \tag{3.19}$$

Condition (C1) implies together with $a_i(k) \sim \sum_{l=1}^{k} \frac{1}{F_i(l)}$ for $k \to \infty$ that

$$\lim_{t\to\infty} o(a_i(t)) \frac{F_i(t)}{t} = 0$$

and hence, replacing t by $a_i^{-1}(t)$ (note: $a_i^{-1}(t) \to \infty$), we get:

$$\lim_{t\to\infty} o(t) \frac{F_i(a_i^{-1}(t))}{a_i^{-1}(t)} = 0$$

Finally,

$$\begin{aligned} \log\left(\frac{a_i^{-1}(t+o(t))}{a_i^{-1}(t)}\right) &= \int_t^{t+o(t)} \frac{d}{dx} \log\left(a_i^{-1}(x)\right) dx = \int_t^{t+o(t)} \frac{\frac{d}{dx} a_i^{-1}(x)}{a_i^{-1}(x)} dx \\ &= \int_t^{t+o(t)} \frac{F_i(a_i^{-1}(x))}{a_i^{-1}(x)} dx \xrightarrow{t\to\infty} 0\,, \end{aligned}$$

which is equivalent to (3.19). □

Theorem 3.24 implies that the market shares in the exponential embedding are asymptotically given by

$$\frac{\Xi_i(t)}{\Xi_1(t)+...+\Xi_A(t)} \sim \frac{a_i^{-1}(t)}{a_1^{-1}(t)+...+a_A^{-1}(t)} \quad \text{for } t \to \infty.$$

Via (2.4) we can now conclude for the discrete-time urn model.

Corollary 3.26. *Let all agents fulfill (C1) and (C2). If the limit*

$$\chi_i(\infty) := \lim_{t\to\infty} \frac{a_i^{-1}(t)}{a_1^{-1}(t)+...+a_A^{-1}(t)} \in [0,1] \tag{3.20}$$

exists for an $i \in [A]$*, then*

$$\chi_i(n) \xrightarrow{n\to\infty} \chi_i(\infty) \quad \textit{almost surely .}$$

If the limit in (3.20) does not exist, then $\chi_i(n)$ *does not converge for* $n \to \infty$*.*
If the limit in (3.20) exists for all $i \in [A]$*, then* $\chi(n) \xrightarrow{n\to\infty} \chi(\infty) \in \Delta_{A-1}$ *almost surely.*

Note that the a_i^{-1} do not depend on N and $\chi(0)$, thus the long time behaviour of market shares $(\chi(n))_{n\in\mathbb{N}}$ in the generalized Pólya urn does not depend on initial conditions if (C1) and (C2) are satisfied. If the limit in (3.20) exists, a market modeled by a Pólya urn under the assumptions of Corollary 3.26 reveals stable and deterministic market shares in the long run and these market shares do not depend on the current market situation and can also take values in $(0,1)$. If the limit $\chi(\infty)$ exists, it is in Δ_{A-1}, since Δ_{A-1} is compact and therefore the laws of $\chi(n)$ form a tight sequence. Corollary 3.26 provides a way to explicitly calculate these long-time market shares.

Example 3.27.

1. If $F_i(k) = \alpha_i k^\beta$ with $\alpha_i > 0, \beta < 1, i \in [A]$, then

   $$a_i^{-1}(t) = (\alpha_i(1-\beta)t+1)^{\frac{1}{1-\beta}}$$

   and hence:

   $$\chi_i(\infty) = \frac{\alpha_i^{\frac{1}{1-\beta}}}{\alpha_1^{\frac{1}{1-\beta}}+...+\alpha_A^{\frac{1}{1-\beta}}} \in (0,1)$$

   Consequently, the impact of the fitness parameters α_i in the long-time limit increases with β, where

   $$\chi_i(\infty) \to \frac{1}{A} \quad \text{for } \beta \to -\infty \quad \text{and} \quad \chi_i(\infty) = \frac{\alpha_i}{\alpha_1+...+\alpha_A} \quad \text{for } \beta = 0\ .$$

   The limiting case $\beta \to 1$ will be discussed later in Proposition 3.30.

2. If $F_i(k) = \alpha_i \log(k+1)$ with $\alpha_i > 0, i \in [A]$, then
$$a_i^{-1}(t) \sim \alpha_i t \log(\alpha_i t) \quad \text{for } t \to \infty$$
and thus:
$$\chi_i(\infty) = \frac{\alpha_i}{\alpha_1 + ... + \alpha_A}$$
Note that this is the same asymptotic market share as if the customers' decisions were independent (with constant feedback functions as for $\beta = 0$ above), so that the strong law of large numbers applies.

It is also possible to find examples where the limit (3.20) does not exist. In the following situation the market share of the agents oscillates with constant amplitude, but increasing period.

Example 3.28. Take $A = 2$ and set
$$a_1^{-1}(t) = t^2 \left(\sin(\log(t)) + 2\right) \quad \text{and} \quad a_2^{-1}(t) = t^2.$$
This corresponds to $F_2(t) = \sqrt{t}$ and $F_1(t) = \left(\frac{d}{dt}(a_1^{-1})^{-1}(t)\right)^{-1}$, which is well defined due to $\frac{d}{dt}a_1^{-1}(t) = t\left(2\sin(\log(t)) + \cos(\log(t)) + 4\right) > 0$. Then Theorem 3.24 implies
$$\frac{\Xi_1(t)}{\Xi_1(t) + \Xi_2(t)} \sim \frac{\sin(\log(t)) + 2}{\sin(\log(t)) + 3} \quad \text{for } t \to \infty$$
and hence $\chi_1(n)$ oscillates between 1/2 and 3/4.

We now add a criterion that ensures the existence of the limit in (3.20).

Corollary 3.29. *Suppose that for an agent i the following tightening of (C1) holds,*
$$\frac{F_i(k)}{k} \sum_{l=1}^{k} \frac{1}{F_i(l)} \xrightarrow{k\to\infty} c \in (0, \infty) \ , \tag{3.21}$$
and that the limits
$$\lim_{k\to\infty} \frac{F_i(k)}{F_j(k)} = c_j \in [0, \infty] \quad \textit{exist for all } j \neq i \ . \tag{3.22}$$
Then the limit in (3.20) exists and
$$\chi_i(\infty) = \left(1 + \sum_{j\neq i} c_j^{-c}\right)^{-1} .$$
In particular, $\mathbb{P}(wMon_i(\chi(0), N)) = 1$ if and only if all c_j are infinity, otherwise $\mathbb{P}(wMon_i(\chi(0), N)) = 0$. If all c_j are one, then Condition (3.21) can be replaced by (C1) and $\chi_j(\infty) = 1/A$ for all $j \in [A]$.

Proof. Recall that $a_i(t) = \int_1^t \frac{dx}{F_i(x)}$ is strictly increasing. For a fixed $j \neq i$ we show that $\Xi_j(t)/\Xi_i(t)$ converges to c_j^{-c}. First, we assume $0 < c_j < \infty$, such that agents i and j fulfill (C1) and (C2). (3.22) implies via the Theorem of de l'Hôspital $a_j(t)/a_i(t) \to c_j$ for $t \to \infty$ and consequently $a_j^{-1}(t) = a_i^{-1}(\delta(t)t)$ for a function δ with $\delta(t) \xrightarrow{t\to\infty} 1/c_j < \infty$. In combination with Theorem 3.24 it remains to show that $a_i^{-1}(\delta(t)t)/a_i^{-1}(t)$ converges to c_j^{-c} for $t \to \infty$. For this we consider

$$\log\left(\frac{a_i^{-1}(\delta(t)t)}{a_i^{-1}(t)}\right) = \int_t^{\delta(t)t} \frac{d}{dx}\log a_i^{-1}(x)dx = \int_t^{\delta(t)t} \frac{F_i(a_i^{-1}(x))}{a_i^{-1}(x)}dx$$
$$\sim \int_t^{\delta(t)t} \frac{c}{x}dx = c\log(\delta(t))$$

as (3.21) implies via time-shift

$$\frac{F_i(a_i^{-1}(t))}{a_i^{-1}(t)} \sim \frac{c}{t} \quad \text{for } t \to \infty.$$

Thus:

$$\frac{a_j^{-1}(t)}{a_i^{-1}(t)} = \frac{a_i^{-1}(\delta(t)t)}{a_i^{-1}(t)} \sim \delta(t)^c \sim c_j^{-c} \quad \text{for } t \to \infty$$

For agents j with $c_j = 0$ the asymptotic market share is for sure bigger than in a situation where F_j is replaced by $CF_i, C > 0$, i.e. $\Xi_j(t)/\Xi_i(t)$ is for $t \to \infty$ larger than any C. Hence, it converges to infinity. Similarly for agents with $c_j = \infty$. □

Note that in the case $c = 1$ (including e.g. feedback functions such as $\log k$, $1/\log k$ or functions converging in $(0, \infty)$) the limit $\chi_i(\infty)$ is equal to the case $F_i(k) = const.$, i.e. draws from the urn are independent and the usual strong law of large numbers applies. So this weak reinforcement does not play any role in the long run.

So far, we did not consider cases near the classical Pólya urn with $F_i(k) = k$, where random limits $\chi_i(\infty)$ are possible. Nevertheless, as Lemma 3.25 does not require (C1), our approach provides some insight into such asymmetric cases as well. The symmetric case with feedback functions close to the classical Pólya urn is treated in Section 3.3.

Proposition 3.30. *Assume that all agents satisfy (C2). In addition, let a fixed agent $i \in [A]$ fulfill*

$$\frac{F_i(k)}{k}\sum_{l=1}^{k}\frac{1}{F_i(l)} \xrightarrow{k\to\infty} \infty, \tag{3.23}$$

but not (M), i.e. $\sum_{k=1}^{\infty}\frac{1}{F_i(k)} = \infty$. *If*

$$\limsup_{k\to\infty}\frac{F_j(k)}{F_i(k)} < 1 \quad \textit{for all agents } j \neq i\ , \tag{3.24}$$

then $\mathbb{P}(wMon_i(\chi(0), N)) = 1$.

Proof. First, note that via exponential embedding, the event $wMon_i(\chi(0), N)$ is equivalent to $\Xi_i(t)/\Xi_j(t) \xrightarrow{t\to\infty} \infty$ for all $j \neq i$. Define

$$\psi(t) \coloneqq \int_0^t \frac{e^x}{F_i(e^x)} dx = \int_1^{e^t} \frac{1}{F_i(x)} dx = a_i(e^t)$$

and thus $a_i^{-1}(t) = e^{\psi^{-1}(t)}$. Assumption (3.24) implies that for any $j \neq i$ there is a constant $c < 1$ with $a_i(t) \leq c a_j(t)$ for large enough t and consequently $a_j^{-1}(t) \leq a_i^{-1}(ct)$. Lemma 3.25 (applicable since (C2) is satisfied) states that $\Xi_i(t) = a_i^{-1}(t+o(t))$ and $\Xi_j(t) = a_j^{-1}(t+o(t)) \leq a_i^{-1}(ct + o(t))$ almost surely. Thus it remains to show that

$$\frac{a_i^{-1}(t + o(t))}{a_i^{-1}(ct + o(t))} \xrightarrow{t\to\infty} \infty,$$

which is equivalent to $\psi^{-1}(t + o(t)) - \psi^{-1}(ct + o(t)) \xrightarrow{t\to\infty} \infty$. It is sufficient that

$$t \cdot \frac{d}{dt}\psi^{-1}(t) = t \cdot \frac{F_i\left(e^{\psi^{-1}(t)}\right)}{e^{\psi^{-1}(t)}} = t\frac{F_i\left(a_i^{-1}(t)\right)}{a_i^{-1}(t)} \xrightarrow{t\to\infty} \infty,$$

which follows since $a_i^{-1}(t) \xrightarrow{t\to\infty} \infty$ and Assumption (3.23) is equivalent to $\frac{F_i(t)}{t} a_i(t) \xrightarrow{t\to\infty} \infty$. □

Condition (3.23) includes feedback functions of the form $F_i(k) = \alpha_i k (\log k)^\beta$ for all $\beta \leq 1$, including the linear case $F_i(k) = \alpha_i k$ for $\beta = 0$. If in addition (3.24) holds, i.e. $\alpha_i > \alpha_j$ for an agent i and all $j \neq i$, then we have an almost sure weak monopoly for agent i. This is consistent with the strong monopoly for $\beta > 1$ as described in Example 3.18. Note that the weak monopoly in Proposition 3.30 is almost sure even for finite N, in contrast to the results on strong monopoly derived in Section 3.1, where the strong monopolist is random and can only be predicted in the limit $N \to \infty$.

On the other hand, Condition (C1) includes sub-linear feedback functions of the form $F_i(k) = \alpha_i k^\beta$ with $\beta < 1$, which have positive long-time market shares for all agents as discussed in Example 3.27.

Exponentially decreasing feedback functions were not taken into account so far as they do not fulfill (C2). Since such cases do not seem to be of great importance for the mentioned interpretations of the model, we are content with an example, which we discuss in Appendix A using the method of stochastic approximation.

3.3 Feedback functions close to the classical Pólya urn

We know from Theorem 2.2 that a generalized Pólya urn reveals strong monopoly if and only if at least one feedback function grows significantly faster than linear, i.e. fulfills (M). As described in Section 3.2, linear feedback functions imply random long-time market shares, whereas a deterministic limit occurs for feedback functions growing significantly slower than linear, i.e. those fulfilling (C1). Nevertheless, some feedback functions that are close to linear (like $F_i(k) = k(\log k)^\beta$, $\beta \neq 0$) are not covered by our results so far. To our knowledge, the literature does not provide results on the long-time behaviour of a generalized Pólya urn with almost linear feedback. For instance, if $F_i(k) = kL(k)$ for a slowly varying function L, then Theorem 2.5 does not determine the long-time limit, since $\lim_{N\to\infty} p(N,x) = x$ for all $x \in \Delta_{A-1}$. We approach this question exploiting general results on birth processes, which require that F_i does not fulfill (M) but inverted squares are summable, i.e.

$$\sum_{k=1}^{\infty} \frac{1}{F_i(k)} = \infty \quad \text{and} \quad \sigma_i^2 := \sum_{k=X_i(0)}^{\infty} \frac{1}{F_i(k)^2} < \infty. \tag{3.25}$$

Recall the exponential embedding from Section 2.1 and notations introduced therein. For this section, it is convenient to adapt previous definitions using

$$a_i(t) := \int_{X_i(0)}^{X_i(0)+t} \frac{dx}{F_i(x)} \quad \text{and} \quad S_i(k) := \sum_{l=X_i(0)}^{X_i(0)+k} \tau_i(l), \tag{3.26}$$

and to extend F_i on $(0,\infty)$ by a right-continuous step function. The key to the desired results is provided by the following result in [11].

Theorem 3.31. *[11, Theorem 3.3', Theorem 3.4, Lemma 3.1] Assume that F_i fulfills (3.25). Then $t - a_i(\Xi_i(t))$ and $S_i(k) - a_i(k)$ converge almost surely for $t \to \infty$ resp. $k \to \infty$ to the same random variable $U_i \in \mathbb{R}$. Moreover, σ_i^2 is the variance of U_i.*

We can now apply this general result in our situation.

Corollary 3.32. *Assume that F_i fulfills (3.25). Then:*

1. *If $\lim_{k\to\infty} \frac{F_i(k)}{k} = 0$, then*

$$\frac{\Xi_i(t)}{a_i^{-1}(t)} \xrightarrow{t\to\infty} 1 \quad \textit{almost surely.}$$

2. *If $\lim_{k\to\infty} \frac{F_i(k)}{k} = c \in (0,\infty)$, then*

$$\frac{\Xi_i(t)}{a_i^{-1}(t)} \xrightarrow{t\to\infty} e^{-cU_i} \quad \textit{almost surely.}$$

3. *If* $\lim_{k\to\infty} \frac{F_i(k)}{k} = \infty$, *then*

$$\frac{\Xi_i(t)}{a_i^{-1}(t)} = \exp\left(\int_t^{t-U_i-o(1)} h_i(s)ds\right)$$

for a (deterministic) function h_i *with* $\lim_{s\to\infty} h_i(s) = \infty$.

Proof. Theorem 3.31 implies $t - a_i(\Xi_i(t)) = U_i + o(1)$ and hence

$$\Xi_i(t) = a_i^{-1}(t - U_i + o(1))$$

Using $\frac{d}{dt}a_i^{-1}(t) = F_i(a_i^{-1}(t) + X_i(0))$ in the logarithmic derivative yields

$$a_i^{-1}(t) = \exp\left(\int_0^t \frac{d}{ds}\log\left(a_i^{-1}(t)\right)ds\right) = \exp\left(\int_0^t \frac{F_i(a_i^{-1}(s) + X_i(0))}{a_i^{-1}(s)}ds\right) \tag{3.27}$$

and consequently

$$\frac{\Xi_i(t)}{a_i^{-1}(t)} = \frac{a_i^{-1}(t - U_i - o(1))}{a_i^{-1}(t)} = \exp\left(\int_t^{t-U_i-o(1)} \frac{F_i(a_i^{-1}(s) + X_i(0))}{a_i^{-1}(s)}ds\right).$$

Now, be aware that $\lim_{t\to\infty} a_i^{-1}(t) = \infty$ as F_i does not fulfill (M) and that the limit of $F(k + const.)/k$ for $k \to \infty$ is equal to the limit of $F(k)/k$. Then all parts of the corollary follow directly from their assumptions. □

Like in Section 3.2, we can now conclude from the exponential embedding to the evolution of market shares in the Pólya urn via

$$\lim_{n\to\infty} \frac{\chi_i(n)}{\chi_1(n) + \ldots \chi_A(n)} = \lim_{t\to\infty} \frac{\Xi_i(t)}{\Xi_1(t) + \ldots \Xi_A(t)},$$

provided that the limit exists.

We are now particularly interested in cases with equal feedback functions for all agents, since agents with different attractiveness are already covered by Proposition 3.30.

Corollary 3.33. *Assume that all agents have the same feedback function* $F_i \equiv F$ *and that* F *fulfills (3.25). Then for all* $\chi(0) \in \Delta^o_{A-1}$*:*

1. *If* $\lim_{k\to\infty} \frac{F(k)}{k} = 0$, *then*

$$\chi_i(n) \xrightarrow{n\to\infty} \frac{1}{A} \quad \textit{almost surely, for all } i \in [A]\ .$$

2. *If* $\lim_{k\to\infty} \frac{F(k)}{k} = c \in (0,\infty)$, *then the limit* $\chi(\infty) = \lim_{n\to\infty} \chi(n)$ *exists almost surely and has a non-degenerate Dirichlet distribution on* Δ_{A-1}.

3. *If* $\lim_{k\to\infty} \frac{F(k)}{k} = \infty$*, then* $\chi(\infty) = \lim_{n\to\infty} \chi(n)$ *exists almost surely and the process exhibits a random weak monopoly, i.e*

$$\mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i(\chi(0), N)\right) = 1 \quad \textit{such that} \quad \mathbb{P}\left(wMon_i(\chi(0), N)\right) > 0 \textit{ for all } i \in [A].$$

In other words: If the feedback function grows any slower than the identity, then the market shares converge to a deterministic limit as time tends to infinity, and the limit does not depend on the initial condition. If the feedback functions grow any faster than the identity, the process exhibits weak monopoly, which is not strong as (M) is necessary in Theorem 2.2. The weak monopoly can been seen in the simulation shown in Figure 2.1 (d). In contrast to the non-symmetric situation of Proposition 3.30, the monopolist is random with probability depending on the initial condition $\chi(0)$.

Proof. Note that the U_i from Theorem 3.31 are independent with distribution depending on $\chi(0)$ and N. In addition, their distribution is continuous as U_i emerges from a sum of independent, centered exponentially distributed random variables. By definition (3.26) we get $a_i(t) = a_j(t + const.) + const.$ with constants depending on the initial conditions and F, and after inversion we have $a_i^{-1}(t) = a_j^{-1}(t + const.) + const.$ for all $i, j \in [A]$. With (3.27) this implies

$$a_i^{-1}(t) = a_j^{-1}(t) \exp\left(\int_t^{t+const.} \frac{F(a_j^{-1}(s) + X_j(0))}{a_j^{-1}(s)} ds\right) + const. \,, \tag{3.28}$$

and note that $a_j^{-1}(t) \to \infty$ in all cases.

1. In this case (3.28) implies that $a_i^{-1}(t) \sim a_j^{-1}(t)$. Then the claim follows directly from Corollary 3.32 via

$$\lim_{n\to\infty} \frac{\chi_i(n)}{\chi_1(n) + \dots \chi_A(n)} = \lim_{t\to\infty} \frac{\Xi_i(t)}{\Xi_1(t) + \dots \Xi_A(t)} = \lim_{t\to\infty} \frac{a_i^{-1}(t)}{a_1^{-1}(t) + \dots a_A^{-1}(t)} = \frac{1}{A} \,.$$

2. Here, again with (3.28), $a_i^{-1}(t)/a_j^{-1}(t)$ converges to a finite, non-zero constant for all $i, j \in [A]$, such that Corollary 3.32 yields

$$\begin{aligned}\lim_{t\to\infty} \frac{\Xi_i(t)}{\Xi_1(t) + \dots \Xi_A(t)} &= \lim_{t\to\infty} \frac{e^{-cU_i} a_i^{-1}(t)}{e^{-cU_1} a_1^{-1}(t) + \dots e^{-cU_A} a_A^{-1}(t)} \\ &= \left(1 + \sum_{j\neq i} e^{c(U_i - U_j)} \lim_{t\to\infty} \frac{a_j^{-1}(t)}{a_i^{-1}(t)}\right)^{-1} .\end{aligned}$$

Hence, $\lim_{n\to\infty} \chi(n)$ exists almost surely and has a continuous distribution, which is of Dirichlet type according to Proposition 2.3 and [78].

3. Due to Lemma 3.34, we can assume $X(0) = (1, \ldots, 1)$, so that $a_i = a_j$ and $h_i = h_j$. Then:

$$\lim_{t\to\infty} \frac{\Xi_i(t)}{\Xi_1(t) + \ldots \Xi_A(t)} = \lim_{t\to\infty} \left(1 + \sum_{j\neq i} \frac{\Xi_j(t)}{\Xi_i(t)}\right)^{-1}$$

$$= \lim_{t\to\infty} \left(1 + \sum_{j\neq i} \exp\left(\int_t^{t-U_j-o(1)} h_j(s)ds - \int_t^{t-U_i-o(1)} h_i(s)ds\right)\right)^{-1}$$

$$= \lim_{t\to\infty} \left(1 + \sum_{j\neq i} \exp\left(\int_{t-U_i-o(1)}^{t-U_j-o(1)} h_i(s)ds\right)\right)^{-1} = \begin{cases} 1 & \text{if } U_i < U_j \text{ for all } j \neq i \\ 0 & \text{else} \end{cases}$$

Recall that $\lim_{s\to\infty} h_i(s) = \infty$. The claim then follows as the U_i are unbounded with continuous distribution (see Lemma 3.37). □

Lemma 3.34. *For all choices of $F_1, \ldots, F_A$, we have*

$$\mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i\left(\left(\frac{1}{A}, \ldots, \frac{1}{A}\right), A\right)\right) = 1$$

$$\Leftrightarrow \quad \mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i(\chi(0), N)\right) = 1 \textit{ for all } \chi(0) \in \Delta^o_{A-1},\ N \in \mathbb{N}.$$

Proof. The implication $\Leftarrow$ is trivial. Thus, assume that the process $X(n)$ starts in $X(0) = (1, \ldots, 1)$ and that $\mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i\left(\left(\frac{1}{A}, \ldots, \frac{1}{A}\right), A\right)\right) = 1$. Moreover, take any $x \in \Delta^o_{A-1}$, $M \in \mathbb{N}$. Then the claim follows directly from the Markov property:

$$1 = \mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i\left(\left(\frac{1}{A}, \ldots, \frac{1}{A}\right), A\right) \mid X(M-A) = Mx\right) = \mathbb{P}\left(\bigcup_{i=1}^{A} wMon_i(x, M)\right),$$

since $\mathbb{P}(X(M-A) = Mx) > 0$ □

The following example presents a class of feedback functions for which four different regimes are possible.

Example 3.35. Let $F_i(k) = k(\log k)^\beta$ for all $i \in [A]$ and $\beta \in \mathbb{R}$. Depending on β, four different regimes occur for $n \to \infty$:

1. For $\beta < 0$, $\chi_i(n)$ for each agent i converges almost surely to $\frac{1}{A}$ independently of $\chi(0)$.
2. For $\beta = 0$, the market shares $\chi(n)$ converge almost surely to a random limit $\chi(\infty) \in \Delta_{A-1}$, which is not a corner point and has a Dirichlet distribution with parameter depending on the initial condition $\chi(0)$.

3. For $\beta \in (0,1]$, the process exhibits a weak monopoly, which is not strong, i.e. all agents win in infinitely many steps, but the market share of one agent converges to one. The monopolist is random, and the distribution of $\chi(\infty)$ on the corner points of Δ_{A-1} depends on $\chi(0)$.

4. For $\beta > 1$, there is a strong monopoly. The monopolist is random and the distribution of $\chi(\infty)$ on the corner points of Δ_{A-1} depends on the initial condition $\chi(0)$ as well.

The remainder of this section is dedicated to the question how fast $\chi(n)$ converges for $n \to \infty$ if the feedback functions are close to the identity. According to Theorem 3.31, we have $t_n - a_i(X_i(n)) \xrightarrow{n\to\infty} U_i$ by definition of the exponential embedding with jump times t_n. If $\lim_{k\to\infty} F(k)/k = \infty$, this convergence can be specified by replacing t_n by a deterministic function and by computing the distribution of U_i.

Theorem 3.36. *Assume that $F_i \equiv F$ does not fulfill (M) and that $F(k)/k \xrightarrow{k\to\infty} \infty$ holds. Then there exist independent random variables $U_1, \dots, U_A$ such that*

$$(a_i(n) - a_i(X_i(n)))_{i\in[A]} \xrightarrow{n\to\infty} \left(U_i - \sum_{k=1}^{X_i(0)-1} \frac{1}{F(k)} - \min_{j\in[A]} \left(U_j - \sum_{k=1}^{X_j(0)-1} \frac{1}{F(k)} \right) \right)_{i\in[A]}$$

almost surely. Moreover, the cumulant generating function (CGF) of each U_i is given by

$$\lambda \mapsto \log\left(\mathbb{E}e^{\lambda U_i}\right) = \sum_{l=2}^{\infty} \frac{\lambda^l}{l} \sum_{k=X_i(0)}^{\infty} \frac{1}{F(k)^l} \tag{3.29}$$

and the radius of convergence is $\min_{k\geq X_i(0)} F(k) > 0$.

In particular, there is exactly one agent, namely the weak monopolist, such that the limit of $a_i(n) - a_i(X_i(n))$ is zero. For the proof, we characterize the distribution of U_i by computing its CGF. For that, we exploit that U_i is also the limit of $S_i(k) - a_i(k)$ for $k \to \infty$ according to Theorem 3.31.

Lemma 3.37. *Assume that $F_i \equiv F$ fulfills (3.25). Then the CGF of U_i is given by (3.29) and the radius of convergence is $\min_{k\geq X_i(0)} F(k)$.*

Proof. The CGF of the limit $U_i = \lim_{k\to\infty} S_i(k) - a_i(k) = \lim_{k\to\infty} \sum_{l=X_i(0)}^{k} \left(\tau_i(l) - \frac{1}{F(l)}\right)$ is the pointwise limit of the CGFs:

$$\begin{aligned}
\log\left(\mathbb{E}e^{\lambda U_i}\right) &= \lim_{k\to\infty} \log\left(\mathbb{E}\exp\left(\lambda \sum_{l=X_i(0)}^{k} \left(\tau_i(l) - \frac{1}{F(l)}\right)\right)\right) \\
&= \log\left(\prod_{k=X_i(0)}^{\infty} e^{-\lambda/F(k)} \mathbb{E}e^{\lambda\tau_i(k)}\right) = \sum_{k=X_i(0)}^{\infty} \left[\log\left(\mathbb{E}e^{\lambda\tau_i(k)}\right) - \frac{\lambda}{F(k)}\right] \\
&= \sum_{k=X_i(0)}^{\infty} \left[\log\left(\frac{F(k)}{F(k)-\lambda}\right) - \frac{\lambda}{F(k)}\right] = -\sum_{k=X_i(0)}^{\infty} \left[\log\left(1 - \frac{\lambda}{F(k)}\right) + \frac{\lambda}{F(k)}\right]
\end{aligned}$$

We now use the series representation of $x \mapsto \log(1+x)$ and change the order of summation due to absolut convergence:

$$\log\left(\mathbb{E}e^{\lambda U_i}\right) = \sum_{k=X_i(0)}^{\infty}\left[\sum_{l=1}^{\infty}\frac{1}{l}\left(\frac{\lambda}{F(k)}\right)^l - \frac{\lambda}{F(k)}\right] = \sum_{k=X_i(0)}^{\infty}\sum_{l=2}^{\infty}\frac{1}{l}\left(\frac{\lambda}{F(k)}\right)^l$$
$$= \sum_{l=2}^{\infty}\frac{\lambda^l}{l}\sum_{k=X_i(0)}^{\infty}\frac{1}{F(k)^l}$$

Note that $\sum_{k=X_i(0)}^{\infty}\frac{1}{F(k)^l} < \infty$ for $l \geq 2$. Now, define $M := \{k \geq X_i(0) \colon F(k) = \min_{l \geq X_i(0)} F(l)\}$ and let $k_0 \in M$. The radius of convergence of the power series representation of the CGF is given by

$$\liminf_{l\to\infty}\left(\frac{1}{l}\sum_{k=X_i(0)}^{\infty}\frac{1}{F(k)^l}\right)^{-1/l} = F(k_0)\liminf_{l\to\infty}\left(\#M + \sum_{k\geq X_i(0),\, k\notin M}\frac{F(k_0)^l}{F(k)^l}\right)^{-1/l} = F(k_0)$$

since $\sum_{k\geq X_i(0),\, k\notin M}\frac{F(k_0)^l}{F(k)^l} \xrightarrow{l\to\infty} 0$ and $\#M < \infty$ if $F(k)/k \xrightarrow{k\to\infty} \infty$. □

In particular, $\mathbb{E}U_i = 0$ since the first term in the series is λ^2, and the l-th cumulant of U_i is $(l-1)!\sum_{k=X_i(0)}^{\infty}\frac{1}{F(k)^l}$ for $l \geq 2$. For the proof of Theorem 3.36, it remains to show that $t_n - a_i(n)$ converges as desired.

Lemma 3.38. *In the situation of Theorem 3.36 we have*

$$t_n - a_i(n) \xrightarrow{n\to\infty} \min_{j\in[A]}(U_j - c_{i,j}),$$

where $c_{i,j} := \sum_{k=1}^{X_j(0)-1}\frac{1}{F(k)} - \sum_{k=1}^{X_i(0)-1}\frac{1}{F(k)}$ *for* $i, j \in [A]$.

Proof. By definition of t_n and $a_i(n)$ and by Theorem 3.31, we have

$$t_n - a_i(n) \leq \min_{j\in[A]} S_j(n) - a_i(n)$$
$$= \min_{j\in[A]}\left(S_j(n) - a_j(n) + a_j(n) - a_i(n)\right) \xrightarrow{n\to\infty} \min_{j\in[A]}(U_j - c_{i,j})$$

as $a_j(n) - a_i(n) \xrightarrow{n\to\infty} c_{i,j}$. Furthermore,

$$t_n - a_i(n) \geq \min_{j\in[A]} S_j(n/A) - a_i(n)$$
$$= \min_{j\in[A]}\left(S_j(n/A) - a_j(n/A) + a_j(n/A) - a_j(n) + a_j(n) - a_i(n)\right) \xrightarrow{n\to\infty} \min_{j\in[A]}(U_j - c_{i,j}).$$

This holds because

$$a_j(n) - a_j(n/A) = \int_{X_j(0)+n/A}^{X_j(0)+n}\frac{1}{F(s)}ds = \left(n - \frac{n}{A}\right)\frac{1}{F(m_n)} \xrightarrow{n\to\infty} 0$$

for a mean value $m_n \in (X_j(0)+n/A,\, X_j(0)+n)$ using $F(k)/k \xrightarrow{k\to\infty} \infty$. □

According to Theorem 3.36, $X_i(n)$ is asymptotically well described by

$$X_i(n) \approx a_i^{-1}\left(a_i(n) - \tilde{U}_i + \min_{j\in[A]} \tilde{U}_j\right),$$

where $\tilde{U}_j := U_j - \sum_{k=1}^{X_j(0)-1} \frac{1}{F(k)}$ for all $j \in [A]$. Now, consider two distinct agents i, j and assume for simplicity of notation that $X_i(0) = X_j(0)$, such that $a_i(k - X_i(0)) = a_j(k - X_j(0)) =: a(k)$. Then Theorem 3.36 states that

$$a(X_i(n)) - a(X_j(n)) \xrightarrow{n\to\infty} U_j - U_i \quad \text{almost surely.}$$

Moreover, the CGF of $U_j - U_i$ is the sum of the CGFs of U_j and U_i due to independence. Hence, $\mathbb{E}e^{\lambda(U_j-U_i)}$ is finite if and only if $|\lambda| < \min_{k\geq X_i(0)} F(k)$. Thus, the distribution of $U_j - U_i$ has exponential tails, and these findings can be used as follows.

Example 3.39. Let $F_i(k) \equiv F(k) = k\log(k)$ and $X_i(0) = X_j(0) = 2$ for two agents $i, j \in [A]$, so that $a_i(t) = a_j(t) = \log\log(t+2) - \log\log 2$. Then the continuous mapping theorem yields

$$\frac{\log X_i(n)}{\log X_j(n)} \xrightarrow{n\to\infty} e^{U_j-U_i} \quad \text{almost surely,}$$

where $e^{U_j-U_i}$ has a power-law distribution due to the explanations above. Remarkably, the log-ratios $\frac{\log X_i(n)}{\log X_j(n)}$ and $\frac{\log X_{i'}(n)}{\log X_{j'}(n)}$ are asymptotically also independent for distinct pairs of agents $(i, j), (i', j')$.

An important application of Theorem 3.36 is its implication for the **rate of convergence**. In fact, the convergence of the process of market shares $\chi(n)$ to a corner of the simplex can be considered as logarithmically slow.

Corollary 3.40. *Assume that $F_i = F$ and $L(k) := F(k)/k$ is non-decreasing, but (M) does not hold. Then there is a random constant $c > 0$ such that*

$$\chi_i(n) \geq e^{-cL(n)} \quad \textit{for all } n \geq 1 \textit{ and } i \in [A].$$

Proof. Since the limit in Theorem 3.36 is finite, there is a constant $c > 0$ such that

$$c \geq \int_{X_i(n)}^{n} \frac{1}{F(s)} ds = \int_{\chi_i(n)}^{1} \frac{1}{sL(ns)} ds \geq \frac{1}{L(n)} \int_{\chi_i(n)}^{1} \frac{1}{s} ds = -\frac{\log(\chi_i(n))}{L(n)},$$

which implies the claim. □

In particular, $\chi_i(n)$ converges to zero slower than any polynomial when $\lim_{n\to\infty} L(n)/\log(n) = 0$. The following example discusses that bound in a generic situation.

Example 3.41. Let $F_i(k) \equiv F(k) = k(\log k)^\beta$ for $\beta \geq 0$. For $\beta = 0$, the lower bound $e^{-cL(n)}$ is constant since $\chi_i(n)$ does converge to a non-zero limit. For $\beta \in (0,1)$, the bound converges to zero slower than any polynomial, whereas it is of order n^{-c} for $\beta = 1$. Note that c is random and unbounded. Finally for $\beta > 1$, the process reveals strong monopoly such that $\chi(n)$ converges to a corner of the simplex at rate $1/n$. In that specific case for $\beta \in (0,1)$, we can also derive an upper bound for $\chi_i(0)$, provided that agent i is not the monopolist. Since the limit in Theorem 3.36 is non-zero and $a_i(t) \sim \frac{1}{1-\beta}(\log t)^{1-\beta}$, there is a positive constant such that for large enough n

$$0 < const. \leq (\log n)^{1-\beta} - (\log(X_i(n)))^{1-\beta}$$

holds and consequently

$$\log(X_i(n)) \leq \Big((\log n)^{1-\beta} - const.\Big)^{\frac{1}{1-\beta}} .$$

Defining $\epsilon(n) = \frac{1}{n}\left(n^{1-\beta} - const.\right)^{\frac{1}{1-\beta}}$ yields:

$$X_i(n) \leq e^{(\log n)\epsilon(\log n)} \quad \Leftrightarrow \quad \frac{X_i(n)}{n} \leq e^{(\log n)\epsilon(\log n) - \log n} = e^{-(\log n)(1-\epsilon(\log n))}$$

Note that $1-\epsilon(n) > 0$ converges to zero at rate $1/n^{1-\beta}$, so that we finally get the following estimate:

$$\chi_i(n) \leq const. e^{-const.(\log n)^\beta}$$

A similar argument is possible for $\beta = 1$. Thus, the bound in Corollary 3.40 can be considered as sharp.

If the second part of (3.25) is not fulfilled, i.e. $\sigma_i^2 = \infty$, then $\frac{S_i(k)-a_i(k)}{\sum_{l=1}^k \frac{1}{F_i(l)^2}}$ fulfills the Lindeberg condition. Hence, Theorem 3.31 and its implications are wrong if we drop the condition $\sigma_i^2 < \infty$. As described in Section 2.2, the leading agent changes infinitely often in time if $\sigma_i^2 = \infty$ and $F_1 = \ldots = F_A$. Theorem 2.11 provides a central limit theorem for this situation.

Another remarkable property is the following: The proof of part 3 of Corollary 3.33 reveals that $\frac{X_i(n)}{X_j(n)} \to 0$ or ∞ for $n \to \infty$ for all $i \neq j$. This corresponds to a hierarchical structure of asymptotic market shares consistent with weak monopoly and the consistency property in Proposition 2.3, such that within each subset of agents a weak monopolist has full relative market share. Such hierarchical structures are often observed at phase transition points, in our case the transition between strong monopoly and deterministic limit shares.

3.4 Overview of the main features of the model

One important purpose of this chapter is to provide a comprehensive approach and a complete picture for the asymptotics of the generalized Pólya urn model, which applies for a large class of feedback functions. This allows us to fully characterize the emergence of monopoly in a transition from sub-linear to super-linear feedback, where the system exhibits interesting behaviour including hierarchical states and weak monopoly. Let us summarize the main results for symmetric feedback $F_1 = \ldots = F_A = F$, although most of the results in this chapter do also hold in asymmetric situations:

1. If F satisfies (M), then the process exhibits strong monopoly (Theorem 2.2 above). The monopolist is random, but can be predicted with high probability for large initial values, such that the space Δ_{A-1} can be dissected into explicitly computable attraction domains (Theorem 3.5).

2. If F does not satisfy (M), but $\lim_{k\to\infty} \frac{F(k)}{k} = \infty$ still holds, then the process exhibits weak monopoly with a random monopolist (Corollary 3.33) and hierarchical states with weak monopoly among the losing agents (Proposition 2.3).

3. If $\lim_{k\to\infty} \frac{F(k)}{k} \in (0, \infty)$, then $\chi(\infty)$ exists almost surely and has a non-degenerate Dirichlet distribution. This includes the classical Pólya urn (Corollary 3.33).

4. If F is sub-linear, then $\chi(\infty)$ exists almost surely and is deterministic with limit given by Corollary 3.26 (under mild, but necessary technical assumptions).

These regimes react differently to unequal fitness of agents. For $F_i(k) = \alpha_i F(k)$ with $\alpha_i > 0$ distinct, we have shown the following properties:

1. If F satisfies (M), then the process still exhibits random strong monopoly with well-defined attraction domains, which continuously depend on α_i for polynomially increasing F. For exponentially increasing F, these domains do not depend on α_i (Theorem 3.14, Corollary 3.8). For F close to the identity, the dependence of the attraction domains on α_i is discontinuous as the fittest agent is almost surely the monopolist in the limit $N \to \infty$ (Example 3.18).

2./3. If F does not satisfy (M), but $\lim_{k\to\infty} \frac{F(k)}{k} \in (0, \infty]$, then the agent with the largest fitness α_i is a deterministic weak monopolist (Proposition 3.30) and we have hierarchical states (Proposition 2.3).

4. If F is sub-linear, then $\chi(\infty)$ still exists and is deterministic. The dependence of $\chi(\infty)$ on α_i can be either continuous (Corollary 3.26) or discontinuous (Proposition 3.30). For exponentially decreasing F, there is no dependence on α_i (Appendix A).

As a conclusion, Pólya's urn with general feedback turns out to be an extremely versatile, which is why it finds various applications in applied mathematical modelling, e.g. in Chapter 6 of this work.

Chapter 4

Functional Limit Theorems

So far, our investigations focused on the analysis of the long-time behaviour of a generalized Pólya urn. This chapter examines the whole dynamics of the process of market shares $(\chi(n))_n$ in the limit for large initial market size N. Our approach is based on the concept of stochastic approximation (see e.g. [107, 27, 100]), which traces back to Robbins and Monro [113]. We derive in Section 4.1 a Functional Law of Large Numbers (LLN) for the dynamics, which are asymptotically described by an Ordinary Differential Equation (ODE). Extending this result, we also establish a Functional Central Limit Theorem (CLT) to describe typical dynamic fluctuations by a system of Stochastic Differential Equations (SDE) in Section 4.2.

For this chapter, it is convenient to establish the notation

$$X^{(N)}(n) := \left(X_1^{(N)}(n), ..., X_A^{(N)}(n)\right) := X(n)$$

and

$$\chi^{(N)}(n) := \left(\chi_1^{(N)}(n), ..., \chi_A^{(N)}(n)\right) := \chi(n)$$

to emphasize the dependence on the initial market size N. We assume that $\chi^{(N)}(0)$ is equal for all N up to roundings, and in particular that $\chi^{(N)}(0)$ converges for $N \to \infty$ to $\chi(0)$.

The results of this chapter have been published in [67].

4.1 A Law of Large Numbers for the dynamics

If the feedback function of at least one agent increases exponentially, then the dynamic of our process in the limit $N \to \infty$ is already fully characterized by Theorem 3.7, since only one predictable agent will win all steps of the process. But for polynomial feedback, several agents will win some steps with high probability for large N as discussed in detail in Example 3.9. In the linear and sub-linear case, the probability of total monopoly is even zero. That is why we need to derive a more refined asymptotic result to describe the typical path of the process of shares towards its limit point, the existence of which is ensured by the imposed assumptions. Recall the definition of the transition probabilities $p(k, x)$ in (2.2).

Theorem 4.1. *Define for* $x \in \Delta_{A-1}$

$$G(k,x) = p(k,x) - x \quad \text{and} \quad G(x) = \lim_{k\to\infty} G(k,x), \tag{4.1}$$

where we assume that $G(k,(\cdot))$ *converges for* $k \to \infty$ *uniformly to a Lipschitz-continuous function* G *on an open neighborhood* $D \subset \Delta_{A-1}$ *of the image of the solution* $Z\colon (0,\infty) \to \Delta_{A-1}$ *of the ODE*

$$\frac{d}{dt} Z(t) = \frac{G(Z(t))}{1+t} \quad \text{with} \quad Z(0) = \chi(0). \tag{4.2}$$

Moreover, we define the following sequence of stochastic processes in Δ_{A-1}*:*

$$(Z^{(N)})_N := \Big(Z^{(N)}(t) \colon t \geq 0\Big)_N := \Big(\chi^{(N)}(\lfloor Nt \rfloor) \colon t \geq 0\Big)_N$$

Then: $Z^{(N)}$ *converges to* Z *weakly on the Skorokhod space* $\mathbb{D}([0,\infty), \Delta_{A-1})$.

Similar ODE approximations in the context of generalized Pólya urns have been derived in [16, 15, 118, 17, 27], but they rather focus on the limit $t \to \infty$ instead of $N \to \infty$, which is trivial in our specific model.

Proof. By construction, we have $\|Z^{(N)}(t) - Z^{(N)}(s)\|_\infty \leq \frac{N|t-s|+1}{N}$ for all $t, s \geq 0$, where $\|\cdot\|_\infty$ denotes the supremum norm. This implies by [77, Proposition VI.3.26] that the sequence $(Z^{(N)})_N$ is tight in $\mathbb{D}([0,\infty), \Delta_{A-1})$, with the additional property that all weak limits of converging subsequences are concentrated on the subspace of continuous functions. We now take any converging subsequence and show that the limit solves (4.2). As the solution of (4.2) is unique due to the assumed Lipschitz-continuity of G (Picard–Lindelöf theorem), this implies the claim. For simplicity of notation assume that the subsequence is $(Z^{(N)})_N$ itself. Then we can write the increments as

$$\begin{aligned}
\chi^{(N)}(n+1) - \chi^{(N)}(n) &= \frac{X^{(N)}(n+1)}{N+n+1} - \chi^{(N)}(n) \\
&= \frac{(N+n)\chi^{(N)}(n) + X^{(N)}(n+1) - X^{(N)}(n)}{N+n+1} - \chi^{(N)}(n) \\
&= \frac{1}{N+n+1}\Big(-\chi^{(N)}(n) + X^{(N)}(n+1) - X^{(N)}(n)\Big) \\
&= \frac{1}{N+n+1}\Big(G(N+n, \chi^{(N)}(n)) + \xi^{(N)}(n)\Big)
\end{aligned}$$

with $\xi^{(N)}(n) := X^{(N)}(n+1) - X^{(N)}(n) - G(N+n, \chi^{(N)}(n)) - \chi^{(N)}(n)$. Note that $\xi^{(N)}(n)$ is $\mathcal{F}^{(N)}_{n+1}$-measurable, where $(\mathcal{F}^{(N)}_n)_{n\geq 0}$ is the filtration generated by the process $(\chi^{(N)}(n))_{n\geq 0}$. Furthermore,

$$\mathbb{E}\Big[\xi^{(N)}(m) \mid \mathcal{F}^{(N)}_n\Big] = 0 \quad \text{for } m \geq n$$

since with (4.1) $\mathbb{E}\Big[X^{(N)}(n+1) - X^{(N)}(n) \mid \mathcal{F}^{(N)}_n\Big] = G(N+n, \chi^{(N)}(n)) + \chi^{(N)}(n)$. The $\xi^{(N)}(n)$ are also uncorrelated, as for $m > n$

$$\mathbb{E}\Big[\xi^{(N)}_i(n)\xi^{(N)}_j(m)\Big] = \mathbb{E}\Big[\xi^{(N)}_i(n)\,\mathbb{E}\Big[\xi^{(N)}_j(m) \mid \mathcal{F}^{(N)}_{n+1}\Big]\Big] = 0 \quad \text{for all } i,j \in [A] \, . \tag{4.3}$$

Summing up the increments yields the standard Doob decomposition

$$\chi^{(N)}(n) = \chi^{(N)}(0) + H^{(N)}(n) + M^{(N)}(n)$$

with predictable and martingale part, respectively

$$H^{(N)}(n) := \sum_{k=0}^{n-1} \frac{G(N+k, \chi^{(N)}(k))}{N+k+1} \quad \text{and} \quad M^{(N)}(n) := \sum_{k=0}^{n-1} \frac{1}{N+k+1} \xi^{(N)}(k) \ . \tag{4.4}$$

With uncorrelated and centered increments $(M^{(N)}(n))_{n\geq 0}$ is a centered martingale with respect to the filtration $(\mathcal{F}_n^{(N)})_{n\geq 0}$, thus Doob's inequality yields for any $\epsilon, t > 0$:

$$\begin{aligned} &\mathbb{P}\left(\exists s \leq t : \|M^{(N)}(\lfloor Ns \rfloor)\|_\infty \geq \epsilon\right) \leq \frac{A}{\epsilon^2} \mathbb{E}\left[\|M^{(N)}(\lfloor Nt \rfloor)\|_\infty^2\right] \\ &\leq \frac{A}{\epsilon^2} \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{1}{(N+k+1)^2} \mathbb{E}\left[\|\xi^{(N)}(k)\|_\infty^2\right] \leq \frac{A}{\epsilon^2} \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{1}{(N+k+1)^2} \xrightarrow{t,N\to\infty} 0 \end{aligned} \tag{4.5}$$

since $\|\xi^{(N)}(k)\|_\infty \leq 1$ almost surely by definition. Hence, the sequence $\left(M^{(N)}(Nt) \colon t \geq 0\right)_N$ of stochastic processes converges to zero weakly on $\mathbb{D}([0,\infty), \mathbb{R}^A)$.

Now, we turn to the predictable part $H^{(N)}$. By Skorokhod's representation theorem, we can find a probability space such that the convergence of $(Z^{(N)})_N$ is almost sure. Then for fixed $\omega \in \Omega$ the sequence of processes $(Z^{(N)})_N$ converges with respect to the Skorokhod norm to a process $\hat{Z}$ on Δ_{A-1}. As $\hat{Z}$ is continuous, the convergence is uniform on bounded time intervals. Denote $t_0 \in (0, \infty]$ the stopping time, when $\hat{Z}$ first leaves D. Then for any $t < t_0$ and large enough $N = N(t)$ we have $Z^{(N)}(t) \in D$ and consequently

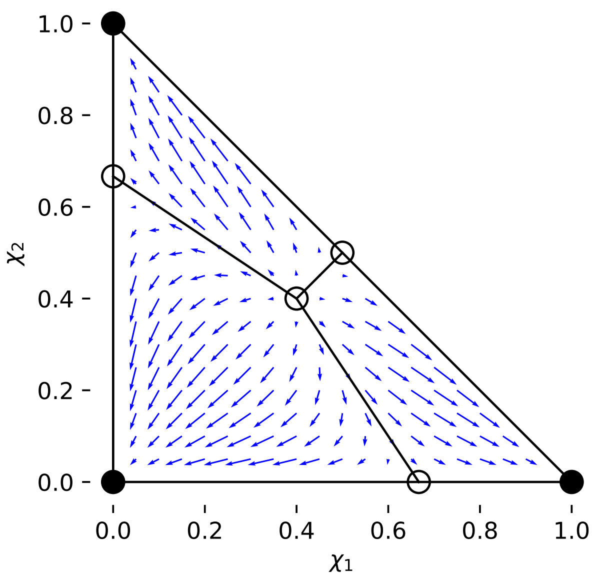
$$\begin{aligned} H^{(N)}(\lfloor Nt \rfloor) = \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{G(N+k, \chi^{(N)}(k))}{N+k+1} &= \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \cdot \frac{G(N+k, \chi^{(N)}(N \cdot \frac{k}{N}))}{1 + \frac{k}{N} + \frac{1}{N}} \\ \xrightarrow{N\to\infty} \int_0^t \frac{G(Z(u))}{1+u} du \end{aligned}$$

as the sequence $\left(u \mapsto \frac{G(N+k, \chi^{(N)}(Nu))}{1+u+\frac{1}{N}}\right)_N$ of functions converges uniformly to $u \mapsto \frac{G(Z(u))}{1+u}$ on bounded time intervals. Thus, we have for $t < t_0$ that $(Z^{(N)})_N$ converges weakly on $\mathbb{D}([0,\infty), \Delta_{A-1})$ to $\hat{Z}(t) = \chi(0) + \int_0^t \frac{G(Z(x))}{1+x} dx$ (which fulfills (4.2)) and by uniqueness of solutions we have $\hat{Z} = Z$ and $t_0 = \infty$. □

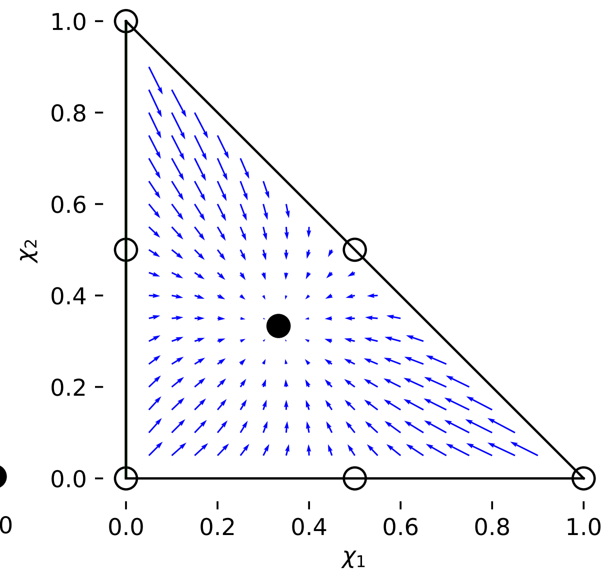
This means, that $(\chi^{(N)}(n))_{n\geq 0}$ is asymptotically deterministic and driven by the vectorfield $(G(x))_{x\in\Delta_{A-1}}$ modulo a time change. Let $Y \colon [0,\infty) \to \Delta_{A-1}$ be the solution of the time-homogeneous ODE

$$\frac{d}{dt} Y(t) = G(Y(t)) \quad \text{with} \quad Y(0) = \chi(0) \ , \tag{4.6}$$

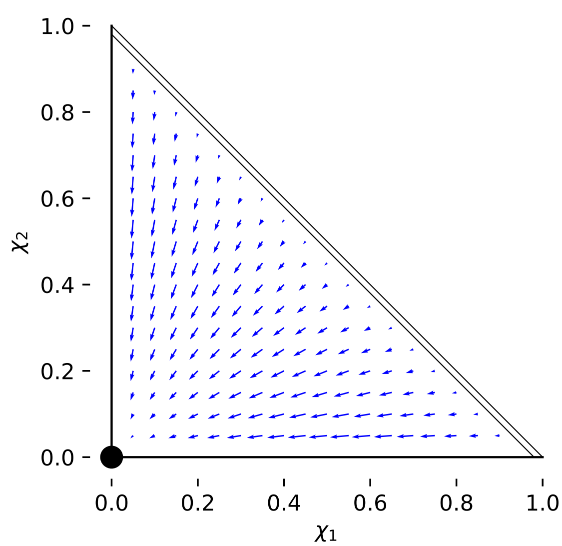
so that $Z(t) = Y(\log(1+t))$. Then for large N the process $(\chi^{(N)}(n))_{n\geq 0}$ is approximately given by $\left(Y\left(\log\left(1 + \frac{n}{N}\right)\right)\right)_{n\geq 0}$. We can use this result e.g. to estimate the number of steps until the process reaches a given neighborhood of its long-time limit for large N.



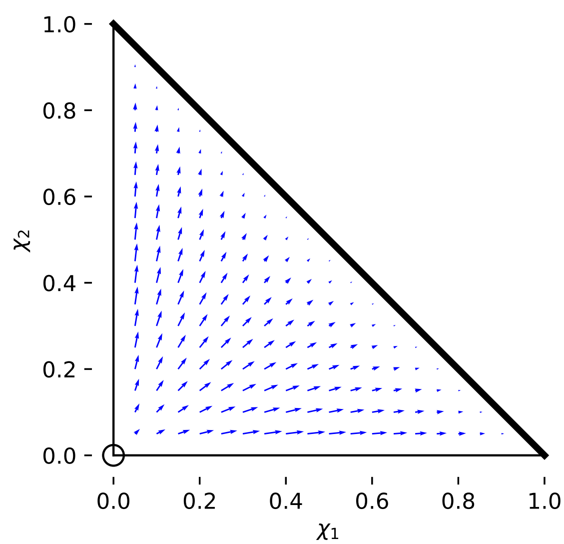
(a) $F_1(k) = F_2(k) = k^2$, $F_3(k) = 2k^2$
$G(x) = \left(\frac{F_i(x_i)}{x_1^2+x_2^2+2x_3^2} - x_i\right)_{i=1,2,3}$, $x \in \Delta_2$



(b) $F_1(k) = F_2(k) = F_3(k) = \sqrt{k}$
$G(x) = \left(\frac{\sqrt{x_i}}{\sqrt{x_1}+\sqrt{x_2}+\sqrt{x_3}} - x_i\right)_{i=1,2,3}$, $x \in \Delta_2$



(c) $F_1(k) = F_2(k) = k$, $F_3(k) = 2k$
$G(x) = \left(\frac{F_i(x_i)}{x_1+x_2+2x_3} - x_i\right)_{i=1,2,3}$, $x \in \Delta_2$



(d) $F_1(k) = F_2(k) = 2k$, $F_3(k) = k$
$G(x) = \left(\frac{F_i(x_i)}{2x_1+2x_2+x_3} - x_i\right)_{i=1,2,3}$, $x \in \Delta_2$

Figure 4.1: The vector field G for different feedback functions and $A = 3$. Here $\bullet$ marks the stable and $\circ$ the unstable fixed points of the dynamics (4.2). In addition, Figure (a) shows the asymptotic attraction domains as derived in Example 3.16.

Corollary 4.2. *In the situation of Theorem 4.1, let* $B \subset \Delta_{A-1}$ *be an open neighborhood of* $\lim_{t\to\infty} Y(t)$ *and define the following last entrance times:*

$$t^* := \sup\{t \geq 0\colon Y(t) \notin B\}$$
$$t_N := \sup\{n \geq 0 : \chi^{(N)}(n) \notin B\}$$

Then we have

$$\frac{t_N}{N} \xrightarrow{N\to\infty} e^{t^*} - 1 \quad \textit{in probability} .$$

This follows directly from the Theorem 4.1 via the continuous mapping theorem.

Another interesting consequence of Theorem 4.1 is the following. In the monopoly case described in Section 3.1, we may start our process in an unstable fixed point $\chi(0)$ of the vector field G. Although we know that the process exhibits strong monopoly, we have $Z(t) \equiv \chi(0)$ for all times $t \geq 0$ in Theorem 4.1. This implies that a linear scaling of time with N is not sufficient to capture the escape from an unstable equilibrium.

Corollary 4.3. *In the situation of Theorem 4.1, let* $G(\chi(0)) = 0$*. For* $\epsilon > 0$ *define the escape time*

$$t_N(\epsilon) := \inf\{n \geq 0\colon \|\chi^{(N)}(n) - \chi^{(N)}(0)\| \geq \epsilon\}.$$

with the convention $\inf \emptyset = \infty$*. Then*

$$\frac{t_N(\epsilon)}{N} \xrightarrow{N\to\infty} \infty \quad \textit{in probability} .$$

Proof. This follows from Theorem 4.1 via

$$\begin{aligned}\mathbb{P}\left(\frac{t_N(\epsilon)}{N} > t\right) &= \mathbb{P}\left(\|Z^{(N)}(s) - \chi(0)\| < \epsilon \text{ for all } s \leq t\right)\\ &= \mathbb{P}\left(\sup_{0\leq s\leq t} \|Z^{(N)}(s) - Z(s)\| < \epsilon\right) \xrightarrow{N\to\infty} 1\end{aligned}$$

for all $t > 0$ since $Z(s) \equiv \chi(0)$. □

Simulations for $F_i(k) = k^\beta$, $\beta > 1$ indicate that the escape from an unstable equilibrium is faster the larger β is. Recall that for super-exponential feedback functions (see Corollary 3.12) the winner of the first step wins in all further steps with high probability if N is large. Hence, it only takes $O(N)$ time to escape from an unstable equilibrium in this case. Nevertheless, this does not pose a contradiction to Corollary 4.3 since the convergence of $G(k, (\cdot))$ to G is not uniform in an unstable equilibrium for super-exponential feedback. Thus, Theorem 4.1 is not applicable and the assumption of uniform convergence cannot be removed.

Theorem 4.1 applies for example for regular varying feedback, i.e. $F_i(k) = \alpha_i k^\beta L(k)$ for all $i \in [A]$ with $\alpha_i > 0$, $\beta \in \mathbb{R}$ and a slowly varying function L. Figure 4.1 shows the dynamics of the process $(\chi(n))_n$ in various generic situations. The fixed points of the dynamics, i.e. the zeros of the vector-field G, are the long-time market-shares of our generalized Pólya urn, but only the stable fixed points are attained with positive probability. Figure (a), (b) and

(c) comply with the properties found in the sections before, i.e. monopoly in the super-linear case and stable, non-zero market-shares in the sub-linear case. Figure (d) underlines that the set of stable fixed points is not necessarily discrete. Note that when $F_i(k) = kL(k)$ for all agents $i \in [A]$ and a slowly varying function L, then the field G is constantly zero, such that all points are fixed points. In particular, this holds for the original Pólya urn, where L is a constant function. If L diverges, then the process exhibits weak monopoly resp. deterministic limits for finite N (see Section 3.3), which is again not captured by Theorem 4.1 as it takes more than $O(N)$ steps to reach the long-time limit.

Moreover, the assumptions of Theorem 4.1 are not always fulfilled for exponential feedback, since G is not continuous. Nevertheless, the dynamics in the limit $N \to \infty$ are already described by Corollary 3.8, which states that all steps are won by the same agents as long as $\chi(0)$ is not on the boundary between the attraction domains. Note that this is consistent with Theorem 4.1, i.e. (4.2) still holds.

Since F_i only depends on X_i and not X_j, $j \neq i$ there are no limit cycles and the dynamics tends to a fixed point of G, as opposed to models discussed in [43].

An alternative interpretation of Theorem 4.1 is the following: We consider an urn process with fixed initial composition $X(0) \in \mathbb{N}^A$. In each time step, we add successively $m \in \mathbb{N}$ balls of weight $\frac{1}{m}$ to the urn following the generalized Pólya urn scheme. According to Theorem 4.1, the composition $X(n) \in (0,\infty)^A$ of that urn after n steps is asymptotically deterministic for $m \to \infty$ and given by $X(n) = Z(n)(N+n)$.

4.2 A Functional Central Limit Theorem for the dynamics

In Section 4.1 we derived a functional LLN for the process of market shares for large initial values, which states that the time-scaled process $Z^{(N)}$ can be well approximated by a deterministic process Z for large N. In order to gain an understanding of the fluctuations around this limit, we prove a corresponding functional CLT in this section. Let us first state our main result. We use the notations introduced in Section 4.1 and establish furthermore the notation

$$p(x) = (p_i(x))_{i\in[A]} = \lim_{k\to\infty} p(k,x) \ , \tag{4.7}$$

for all $x \in \Delta_{A-1}$. Note that the existence of p is equivalent to the existence of G. Denote by

$$T\Delta_{A-1} \coloneqq \left\{ (x_1,\dots,x_A) \in \mathbb{R}^A \colon \sum_{i=1}^{A} x_i = 0 \right\} \tag{4.8}$$

the tangent space of Δ_{A-1}.

Theorem 4.4. *Suppose that*

$$\lim_{k\to\infty} \sqrt{k} \sup_{x\in\Delta_{A-1}} \|G(k,x) - G(x)\| = 0. \tag{4.9}$$

Moreover, let G be continuously differentiable on Δ^o_{A-1}. Then we have

$$\sqrt{N}\left(Z^{(N)}(t) - Z(t)\right)_{t\geq 0} \xrightarrow{N\to\infty} (\tilde{Z}(t))_{t\geq 0} \quad \textit{weakly on } \mathbb{D}([0,\infty), T\Delta_{A-1}),$$

where $\tilde{Z}$ is the solution of the system of SDEs

$$d\tilde{Z}_i(t) = \frac{DG_i(Z(t))}{1+t} \cdot \tilde{Z}(t)dt + \sum_{j\neq i} \frac{\sqrt{p_i(Z(t))p_j(Z(t))}}{1+t} dB_{i,j}(t), \quad i \in [A]. \tag{4.10}$$

Here, DG_i denotes the differential operator of G_i as explained below and $B_{i,j}$ is a standard Brownian motion, which is independent of $B_{k,l}$ if $\{i,j\} \neq \{k,l\}$ and $B_{j,i} = -B_{i,j}$ for $i \neq j$.

The differential operator $DG_i(z)\colon T\Delta_{A-1} \to \mathbb{R}$, $\tilde{z} \mapsto DG_i(z) \cdot \tilde{z}$ for $z \in \Delta^o_{A-1}$ is the product with the gradient $\nabla G_i(z)$, when G is defined on an open neighbourhood of $T\Delta_{A-1}$ in $\mathbb{R}^A$. [27] (and similarly [6]) presents a functional CLT in a general stochastic approximation setting, but they consider rather the limit $t \to \infty$ than $N \to \infty$. [28, 39] recently proved a corresponding result for the classical Pólya urn and [46, 70, 79, 8] for Friedman's urn.

For the proof, we use again the method of stochastic approximation. In the Doob decomposition (4.4), we prove separately a limit theorem for the martingale part $M^{(N)}$ in Subsection 4.2.1 and for the predictable part $H^{(N)}$ in Subsection 4.2.2, which directly imply Theorem 4.4 by summing up both. Note that Theorem 4.7 for the martingale does not use the rather restrictive condition (4.9). Within these subsections, we discuss in detail the properties and interpretation of the diffusion part and the drift part of (4.10).

Figure 4.2 shows the process $Z^{(N)} - Z$ for large N. We can observe that $Z^{(N)}(t) - Z(t)$ is close to zero for large t. Indeed, this complies with formula (4.10).

Proposition 4.5. *In the situation of Theorem 4.4, assume that $Z(\infty) := \lim_{t\to\infty} Z(t)$ exists and that $DG(Z(\infty))$ is a negative definite operator. Then*

$$\tilde{Z}(t) \xrightarrow{t\to\infty} 0 \quad \text{in } L^2 \text{ and almost surely.}$$

Proof. As explained in Subsection 4.2.1, the generator of $\tilde{Z}$ is given by

$$\begin{aligned} L_t f(x) = &\sum_{i=1}^{A} \frac{DG_i(Z(t)) \cdot x}{1+t} \cdot \frac{\partial}{\partial x_i} f(x) + \sum_{i=1}^{A} \frac{p_i(Z(t))(1-p_i(Z(t)))}{2(1+t)^2} \frac{\partial^2}{\partial x_i^2} f(x) \\ &+ \sum_{\substack{i=1\\ i\neq j}}^{A} \frac{p_i(Z(t))p_j(Z(t))}{(1+t)^2} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \end{aligned}$$

for $x = (x_1, \ldots, x_A) \in T\Delta_{A-1}$. Thus, for $f(x) = x_1^2 + \ldots + x_A^2$ we have

$$\begin{aligned} L_t f(x) &= \sum_{i=1}^{A} \frac{DG_i(Z(t)) \cdot x}{1+t} \cdot 2x_i + \sum_{i=1}^{A} \frac{p_i(Z(t))(1-p_i(Z(t)))}{(1+t)^2} \\ &= \frac{2}{1+t} \langle DG(Z(t))x, x\rangle + \frac{b(t)}{(1+t)^2} \end{aligned}$$

for a bounded function $b(t)$. Since $t \mapsto DG(Z(t))$ is continuous and $DG(Z(\infty))$ is negative definite, $DG(Z(t))$ is also negative definite for $t \geq t_0$, when $t_0 > 0$ is large enough. Thus, there is $\lambda > 0$ such that

$$\langle DG(Z(t))x, x\rangle \leq -\lambda \|x\|^2 \tag{4.11}$$

for all $x \in \mathbb{R}^A$ and $t \geq t_0$. In summary, we get

$$L_t f(x) \leq -\frac{2\lambda}{1+t}\|x\|^2 + \frac{b(t)}{(1+t)^2}$$

for $t \geq t_0$. Now, applying Dynkin's formula yields

$$\frac{d}{dt}\mathbb{E}\|\tilde{Z}(t)\|^2 = \mathbb{E}L_t f(\tilde{Z}(t)) \leq -\frac{2\lambda}{1+t}\mathbb{E}\|\tilde{Z}(t)\|^2 + \frac{b(t)}{(1+t)^2}.$$

for $t \geq t_0$. Finally, the claim follows from Grönwall's inequality:

$$\mathbb{E}\|\tilde{Z}(t)\|^2 \leq \left(\int_{t_0}^t \frac{b(s)}{(1+s)^2}ds + \mathbb{E}\|\tilde{Z}(t_0)\|^2\right)\exp\left(\int_{t_0}^t -\frac{2\lambda}{1+s}ds\right) \xrightarrow{t\to\infty} 0$$

For the almost sure convergence, we consider the limiting processes M and H derived in Theorem 4.11 and Theorem 4.12. We fix a realisation $\omega \in \Omega$, such that $m := \lim_{t\to\infty} M(t)(\omega)$ exists. Then we get from (4.15) and the Cauchy-Schwarz inequality that

$$\begin{aligned}
&\frac{d}{dt}\|H(t)+m\|^2 \\
&= \frac{2}{1+t}\big(\langle DG(Z(t))(H(t)+m), H(t)+m\rangle + \langle H(t)+m, DG(Z(t))(M(t)-m)\rangle\big) \\
&\leq \frac{2}{1+t}\big(-\lambda\|H(t)+m\|^2 + \|H(t)+m)\| \cdot \|DG(Z(t))(M(t)-m)\|\big)
\end{aligned}$$

for $t \geq t_0$. Hence

$$\frac{d}{dt}\|H(t)+m\|^2 > 0 \quad \Longrightarrow \quad \frac{\|DG(Z(t))(M(t)-m)\|}{\lambda} > \|H(t)+m\|,$$

which implies $\|H(t)+m\| \xrightarrow{t\to\infty} 0$ as $\|DG(Z(t))(M(t)-m)\| \xrightarrow{t\to\infty} 0$. □

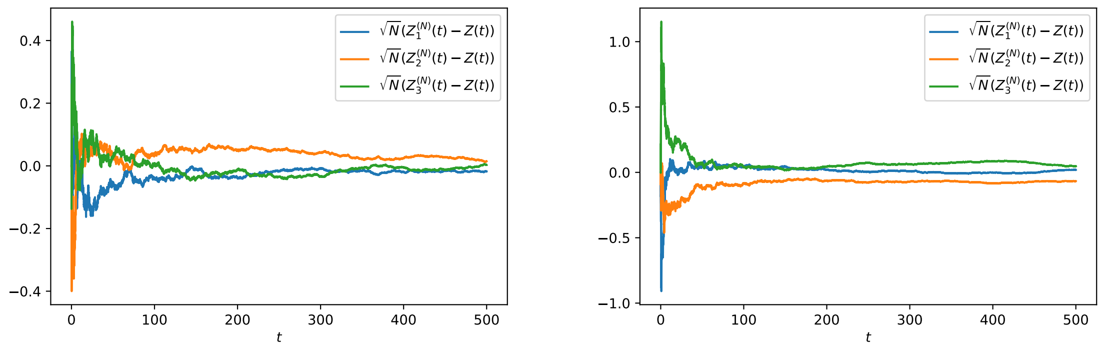
In generic examples one can show that $DG(Z(\infty))$ is indeed negative definite, but it is also possible to find a counterexample.

Example 4.6. Let $F_i(k) = \alpha_i k^\beta$ for $\alpha_i > 0$, $\beta > 0$, such that

$$G_i(x) = \frac{\alpha_i x_i^\beta}{\alpha_1 x_1^\beta + \ldots \alpha_A x_A^\beta} - x_i \quad \text{for all } x \in \Delta_{A-1}\ .$$

Since there is an obvious extension of G to $\mathbb{R}^A$, the operator $DG(x)$ is negative definite if and only if the well-defined differential matrix $\left(\frac{\partial}{\partial x_j}G_i(x)\right)_{i,j=1,\ldots,A}$ is negative definite.

1. Consider the monopoly case $\beta > 1$. Moreover, let $\chi(0)$ be in the attraction domain of agent i, i.e. $Z(t) \xrightarrow{t\to\infty} e^{(i)}$. A simple computation shows $\nabla G_j(e^{(i)}) = (-\delta_{l,j})_{l=1,\ldots,A}$ for all $j \in [A]$, where $\delta_{i,j}$ denotes the Kronecker delta. Hence, $DG(e^{(i)})$ is negative definite.



(a) $F_1(k) = F_2(k) = F_3(k) = \sqrt{k}$, $\chi(0) = \left(\frac{1}{10}, \frac{1}{10}, \frac{4}{5}\right)$

(b) $F_1(k) = F_2(k) = F_3(k) = k^2$, $\chi(0) = \left(\frac{3}{10}, \frac{3}{10}, \frac{2}{5}\right)$

Figure 4.2: The processes $\sqrt{N}\left(Z^{(N)}(t) - Z(t)\right)$ for $A = 3$ and $N = 10.000$.

2. In the monopoly case $\beta > 1$ assume that $\chi(0)$ is the unique unstable fixpoint of the vector field G in Δ^o_{A-1}. Then $Z(\infty) = \chi(0)$ and $DG(Z(\infty))$ is positive definite. Thus, $\mathbb{E}\|\tilde{Z}(t)\|^2 \xrightarrow{t\to\infty} \infty$ follows by similar argumentation.

3. For $\beta = 1$, we have $H(t) \equiv 0$ since $G(x) \equiv 0$. In this case $\tilde{Z}(t)$ does not converge to zero for $t \to \infty$. This is due to the fact that for $\beta = 1$ and large (but finite) N the time-limit $\lim_{n\to\infty} \chi^{(N)}(n)$ is close to $\chi(0)$, but still random. For $\beta \neq 1$, the long-time limit can be predicted precisely for large N (at least with high probability).

4. Now, let $\beta < 1$. For simplicity, assume $\alpha_i = 1$ for all $i \in [A]$, but a similar argument is possible in a non-symmetric situation. Then $Z(\infty) := \lim_{t\to\infty} Z(t) = \left(\frac{1}{A}\right)_{i=1,\dots,A}$. It can be shown that $\nabla G_i(Z(\infty)) = (c\delta_{i,j} + d(1-\delta_{i,j}))_{j=1,\dots,A}$ for some $c < d < 0$, i.e. $DG(Z(\infty))$ is negative definite.

Note that the time-change factor $\frac{1}{1+t}$ in (4.10) does not change the long-time limit of the dynamics, but slows down the rate of convergence. The Grönwall estimate in the proof of Proposition 4.5 implies that $\tilde{Z}(t)$ converges to zero at least at rate $t^{-2\lambda}$, where $\lambda > 0$ satisfies (4.11). For the classical Pòlya urn we have $\lambda = 0$, such that there is no convergence to zero.

As we can see, the first steps of our process are of particular interest. This is investigated in detail in Appendix B for the limiting behaviour of $(\chi(\lfloor N^\beta t \rfloor))_{t\geq 0}$ for $N \to \infty$ and non-linear time scale $\beta \in (0, 1)$.

4.2.1 Convergence of the martingale part

This subsection examines the martingale $M^{(N)} = (M_1^{(N)}, \dots, M_A^{(N)})$ as defined in (4.4). We have already seen in Section 4.1 that $M^{(N)}$ vanishes for $N \to \infty$. Under appropriate scaling, we can yield the following CLT, which accounts for the diffusion part of (4.10). For simplicity we will at first only consider one fixed agent (without loss of generality agent 1) while keeping $A \geq 2$ general.

Theorem 4.7. *We assume that the convergence (4.7) is uniform on an open neighborhood of the image of Z and that p is a Lipschitz continuous function on this neighbourhood. Moreover, denote by $(M_1(t))_{t\geq 0}$ a time-inhomogeneous Markov process with generator*

$$L_s f := \frac{f''}{2(1+s)^2} p_1(Z(s))(1-p_1(Z(s))), \quad s \geq 0$$

and $M_1(0) = 0$. Then

$$\sqrt{N}\left(M_1^{(N)}(\lfloor Nt \rfloor)\right)_{t\geq 0} \xrightarrow{N\to\infty} (M_1(t))_{t\geq 0} \quad \text{weakly on } \mathbb{D}([0,\infty),\mathbb{R}) \ .$$

Alternatively, the inhomogeneous Markov-process M_1 is characterized as the solution of the SDE

$$dM_1(t) = \frac{\sqrt{p_1(Z(t))(1-p_1(Z(t)))}}{1+t} dB(t), \quad M_1(0) = 0,$$

where B denotes a standard Brownian motion and Z is the solution of the ODE (4.2). Thus, M_1 is a time-changed Brownian motion. To be more precise, $M_1(t) = B(\langle M \rangle_t)$, where

$$t \mapsto \langle M_1 \rangle(t) := \int_0^t \frac{p_1(Z(s))(1-p_1(Z(s)))}{(1+s)^2} dt \leq \int_0^t \frac{1}{4(1+s)^2} ds < \frac{1}{4}$$

is the quadratic variation process of M_1. Note that $\langle M_1 \rangle(t)$ is deterministic and increasing in t, and thus $M_1(t)$ converges almost surely for $t \to \infty$ and the limit has a centered Gaussian distribution with variance $\lim_{t\to\infty} \langle M_1 \rangle(t)$.

For the proof of Theorem 4.7, we first show tightness of the sequence $\left(\sqrt{N} M_1^{(N)}(\lfloor Nt \rfloor) \colon t \geq 0\right)_N$ on $\mathbb{D}([0,\infty),\mathbb{R})$ and then prove that the limit of any converging subsequence is a Markov-process with generator $(L_s)_{s>0}$. For later use in Appendix B, we keep the tightness result a bit more general than necessary.

Lemma 4.8. *The sequence of martingales $\left(N^{1-\frac{\beta}{2}} M_1^{(N)}(\lfloor N^\beta t \rfloor) \colon t \geq 0\right)_N$ is tight for all $\beta \in (0,1]$.*

Proof. According to a version of the Aldous criterion in [128, Lemma 3.11], the following two properties are sufficient for the tightness.

1. Stochastic Boundedness: For $C, T > 0$ we have by Doob's inequality and (4.3)

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0<t\leq T} N^{1-\frac{\beta}{2}}\left|M_1^{(N)}(\lfloor N^\beta t\rfloor)\right| > C\right) \leq \frac{N^{2-\beta}}{C^2}\mathbb{E}\left(M_1^{(N)}(\lfloor N^\beta T\rfloor)^2\right)\\
&= \frac{N^{2-\beta}}{C^2}\sum_{k=0}^{\lfloor N^\beta T\rfloor-1}\frac{1}{(N+k+1)^2}\mathbb{E}\left(\xi_1^{(N)}(k)^2\right) \leq \frac{N^{2-\beta}}{C^2}\sum_{k=0}^{\lfloor N^\beta T\rfloor-1}\frac{1}{(N+k+1)^2}\\
&\leq \frac{N^{2-\beta}}{C^2}\int_N^{\lfloor N^\beta T\rfloor+N}\frac{1}{s^2}ds = \frac{N^{2-\beta}}{C^2}\left(\frac{1}{N}-\frac{1}{\lfloor N^\beta T\rfloor+N}\right) \leq \frac{N^{2-\beta}}{C^2}\cdot\frac{\lfloor N^\beta T\rfloor}{N^2}\\
&\leq const.(T)/C^2 \xrightarrow{C\to\infty} 0
\end{aligned}$$

uniformly in N.

2. Similarly, we get for $0 < t \leq T$ and $0 < u \leq \delta$:

$$\begin{aligned}
&\mathbb{E}\left[\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(\lfloor N^\beta(t+u)\rfloor) - N^{1-\frac{\beta}{2}}M_1^{(N)}(\lfloor N^\beta t\rfloor)\right)^2 \left|\mathcal{F}_{\lfloor N^\beta t\rfloor}^{(N)}\right.\right]\\
&\leq N^{2-\beta}\sum_{k=\lfloor N^\beta t\rfloor}^{\lfloor N^\beta(t+u)\rfloor-1}\frac{1}{(N+k+1)^2}\mathbb{E}\left[\xi_1^{(N)}(k)^2\middle|\mathcal{F}_{\lfloor N^\beta t\rfloor}^{(N)}\right] \leq N^{2-\beta}\sum_{k=\lfloor N^\beta t\rfloor}^{\lfloor N^\beta(t+\delta)\rfloor-1}\frac{1}{(N+k+1)^2}\\
&\leq N^{2-\beta}\int_{\lfloor N^\beta t\rfloor+N}^{\lfloor N^\beta(t+\delta)\rfloor+N}\frac{1}{s^2}ds = N^{2-\beta}\left(\frac{1}{\lfloor N^\beta t\rfloor+N}-\frac{1}{\lfloor N^\beta(t+\delta)\rfloor+N}\right)\\
&\leq N^{2-\beta}\cdot\frac{\lfloor N^\beta(t+\delta)\rfloor-\lfloor N^\beta t\rfloor}{N^2} \leq const.(\delta) \xrightarrow{\delta\to 0} 0
\end{aligned}$$

uniformly in N. □

By the definition of tightness and Theorem 4.1, we also get tightness of the joint sequence $(Z^{(N)}, N^{1-\frac{\beta}{2}}M_1^{(N)}(\lfloor N^\beta(\cdot)\rfloor))_N$. Before we turn to the proof of Theorem 4.7, we add another helpful lemma.

Lemma 4.9. *With p as defined in (4.7) and $\beta \in (0,1]$, we have for all smooth test-functions $f : \mathbb{R} \to \mathbb{R}$ with compact support*

$$\begin{aligned}
&\mathbb{E}\left[f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k+1)\right) - f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)\middle|\mathcal{F}_k^{(N)}\right]\\
&= \frac{N^{2-\beta}}{2(N+k+1)^2}f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)p_1(N+k,\chi^{(N)}(k))\left(1-p_1(N+k,\chi^{(N)}(k))\right) + o\left(N^{-\beta}\right)
\end{aligned}$$

as $N \to \infty$.

Proof. Taylor-expansion of f with Lagrange's remainder yields:

$$
\begin{aligned}
&E\left[f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k+1)\right)-f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)|\mathcal{F}_k^{(N)}\right]\\
&=N^{1-\frac{\beta}{2}}f'\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)\mathbb{E}\left[M_1^{(N)}(k+1)-M_1^{(N)}(k)|\mathcal{F}_k^{(N)}\right]\\
&\quad+\frac{N^{2-\beta}}{2}\mathbb{E}\left[f''\left(m^{(N)}(k)\right)\left(M_1^{(N)}(k+1)-M_1^{(N)}(k)\right)^2|\mathcal{F}_k^{(N)}\right]\\
&=\frac{N^{2-\beta}}{2(N+k+1)^2}\left[\mathbb{E}\left[f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)\xi_1^{(N)}(k)^2|\mathcal{F}_k^{(N)}\right]+o(1)\right]\\
&=\frac{N^{2-\beta}}{2(N+k+1)^2}f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)\left(\left(1-p_1(N+k,\chi^{(N)}(k))\right)^2p_1(N+k,\chi^{(N)}(k))\right.\\
&\quad\left.+p_1(N+k,\chi^{(N)}(k))^2\left(1-p_1(N+k,\chi^{(N)}(k))\right)\right)+o\left(N^{-\beta}\right)\\
&=\frac{N^{2-\beta}}{2(N+k+1)^2}f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)p_1(N+k,\chi^{(N)}(k))\left(1-p_1(N+k,\chi^{(N)}(k))\right)+o\left(N^{-\beta}\right)
\end{aligned}
$$

Here, $m^{(N)}(k)$ denotes a (random) intermediate value between $N^{1-\frac{\beta}{2}}M_1^{(N)}(k)$ and $N^{1-\frac{\beta}{2}}M_1^{(N)}(k+1)$. Note that $m^{(N)}(k)-N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\xrightarrow{N\to\infty}0$ at rate $N^{-\frac{\beta}{2}}$. □

Now we are well prepared for the proof of Theorem 4.7.

Proof. We show that for any limit (Z,M_1) of a convergent subsequence of $(Z^{(N)},\sqrt{N}M_1^{(N)}(\lfloor N(\cdot)\rfloor))_N$, M_1 is a Markov process with generator $(L_s)_{s>0}$. For simplicity of notation, assume that the sequence is convergent itself.

Take a smooth test-function $f\colon\mathbb{R}\to\mathbb{R}$ with compact support. Then for each N

$$
f\left(\sqrt{N}M_1^{(N)}(\lfloor Nt\rfloor)\right)-f(0)-\sum_{k=0}^{\lfloor Nt\rfloor-1}\mathbb{E}\left[f\left(\sqrt{N}M_1^{(N)}(k+1)\right)-f\left(\sqrt{N}M_1^{(N)}(k)\right)|\mathcal{F}_k^{(N)}\right],\tag{4.12}
$$

is a martingale in continuous time $t\geq 0$ as $(Z^{(N)},M_1^{(N)})$ is a discrete-time Markov process. The continuous mapping theorem implies that $f\left(\sqrt{N}M_1^{(N)}(\lfloor N(\cdot)\rfloor)\right)$ converges to $f(M_1)$ in

$\mathbb{D}((0,\infty),\mathbb{R})$. Due to Lemma 4.9, the sum converges as follows:

$$\begin{aligned}
&\sum_{k=0}^{\lfloor Nt\rfloor-1} \mathbb{E}\left[f\left(\sqrt{N}M_1^{(N)}(k+1)\right) - f\left(\sqrt{N}M_1^{(N)}(k)\right)|\mathcal{F}_k^{(N)}\right] \\
&= \sum_{k=0}^{\lfloor Nt\rfloor-1} \Bigg[\frac{N}{2(N+k+1)^2} f''\left(\sqrt{N}M_1^{(N)}(k)\right) p_1(N+k,\chi^{(N)}(k)) \left(1-p_1(N+k,\chi^{(N)}(k))\right) \\
&\qquad + o(1/N)\Bigg] \\
&= \sum_{k=0}^{\lfloor Nt\rfloor-1} \frac{1}{2N(1+\frac{k}{N}+\frac{1}{N})^2} f''\left(\sqrt{N}M_1^{(N)}\left(N\frac{k}{N}\right)\right)\Bigg) p_1\left(N+k, Z^{(N)}\left(\frac{k}{N}\right)\right) \\
&\quad \cdot \left(1-p_1\left(N+k, Z^{(N)}\left(\frac{k}{N}\right)\right)\right) + o(1) \\
&\xrightarrow{N\to\infty} \int_0^t \frac{f''(M_1(s))}{2(1+s)^2} p_1(Z(s))\,(1-p_1(Z(s))\,ds = \int_0^t L_s f(M_1(s))ds
\end{aligned}$$

Convergence for $N \to \infty$ holds almost surely on an appropriate probability space by Skorokhod's representation theorem, which implies weak convergence. Summing up, we have that (4.12) converges to

$$f(M_1(t)) - f(0) - \int_0^t L_s f(M_1(s))ds \tag{4.13}$$

for $N \to \infty$. As f and f'' are bounded, the sequence in (4.12) is obviously uniformly integrable in N. Thus, [128, Theorem 5.3] implies that (4.13) is a martingale as well. Moreover, the solution of the martingale problem (4.13) is unique as a time-changed Brownian motion is always the unique solution if its corresponding martingale problem. Hence, M_1 is a time-inhomogeneous Markov-process with generator $(L_s)_{s\geq 0}$. □

Example 4.10.

1. Let $F_i(k) = e^{\beta_i k}$, $\beta_i > 0$, $i \in [A]$ and suppose that $\chi_i(0)\alpha_i > \chi_j(0)\alpha_j$ for an $i \in [A]$ and all $j \neq i$, i.e. $\chi(0)$ is in the attraction domain D_i of agent i. Then $M_1(t) = 0$ almost surely for all $t \geq 0$, since $p(x) = e^{(i)}$ for $x \in D_i$, in particular on the path of Z. This complies with the idea of a total monopoly described in Section 3.1.

2. If $F_i(k) = k$, $i \in [A]$, then $Z(s) \equiv \chi(0)$ for all $s \geq 0$ and $p(x) = x$ for all $x \in \Delta_{A-1}$. Hence, $\langle M_1\rangle(t) = \chi_1(0)(1-\chi_1(0))\left(1-\frac{1}{1+t}\right)$ for all $t \geq 0$. Note that in this case the martingale part $M^{(N)} = \chi^{(N)} - \chi^{(N)}(0)$ encompasses the whole dynamic as $H^{(N)}(t) \equiv 0$ for all $t \geq 0$.

3. Let $F_i(k) = k^\beta$, $\chi_i(0) = \frac{1}{A}$ for all $i \in [A]$ and $\beta > 0$. Since we start in a stable or unstable equilibrium point, we have $Z(t) \equiv \chi(0)$ and hence $\langle M_1\rangle(t) = \frac{A-1}{A^2}\left(1-\frac{1}{1+t}\right)$ for all $t \geq 0$. In particular, M_1 does not depend on β. For $\beta > 1$, M_1 describes the fluctuations around the unstable fixed point, but does not capture the escape from it.

For non-linear, polynomial feedback functions and general initial market shares, the expressions for Z are lengthy or even not explicit. Figure 4.3 shows some realisations of the process M_1. It can be seen that the convergence of $M_1(t)$ for $t \to \infty$ is faster the faster the feedback functions grow. In the monopoly case, the variation of M_1 is small if $\chi(0)$ is already close to zero or one.

So far in this section, we only considered one fixed agent. Nevertheless, one can obtain an extension of Theorem 4.7 for all agents by a completely analogous, but lengthy argument, which we leave to the reader.

Theorem 4.11. *Suppose that the assumptions of Theorem 4.7 are fulfilled. Moreover, denote by $(M(t))_{t\geq 0}$ an A-dimensional time-inhomogeneous Markov process with generator*

$$\tilde{L}_s f(x) := \sum_{i=1}^{A} \frac{p_i(Z(s))(1-p_i(Z(s))}{2(1+s)^2} \frac{\partial^2}{(\partial x_i)^2} f(x) - \sum_{\substack{i,j=1 \\ j\neq i}}^{A} \frac{p_i(Z(s))p_j(Z(s))}{(1+s)^2} \frac{\partial^2}{\partial x_i \partial x_j} f(x)$$

with $x \in \mathbb{R}^A$ and $M(0) = 0$. Then

$$\sqrt{N}\left(M^{(N)}(\lfloor Nt \rfloor)\right)_{t\geq 0} \xrightarrow{N\to\infty} (M(t))_{t\geq 0} \quad \textit{weakly on } \mathbb{D}([0,\infty),\mathbb{R}^A).$$

The specific form of the generator is due to the conditioned covariance matrix of the increments $\xi^{(N)}$, which is for $j \neq i$:

$$\begin{aligned}\mathbb{E}\left[\xi_i^{(N)}(k)\xi_j^{(N)}\middle|\mathcal{F}_k^{(N)}\right] &= -p_i(N+k,\chi^{(N)}(k))\left(1-p_i(N+k,\chi^{(N)}(k))\right)p_j(N+k,\chi^{(N)}(k)) \\ &\quad - p_j(N+k,\chi^{(N)}(k))\,p_i(N+k,\chi^{(N)}(k))\left(1-p_j(N+k,\chi^{(N)}(k))\right) \\ &\quad + \left(1-p_i(N+k,\chi^{(N)}(k))-p_i(N+k,\chi^{(N)}(k))\right)p_i(N+k,\chi^{(N)}(k))\,p_j(N+k,\chi^{(N)}(k)) \\ &= -p_i(N+k,\chi^{(N)}(k))\,p_j(N+k,\chi^{(N)}(k))\end{aligned}$$

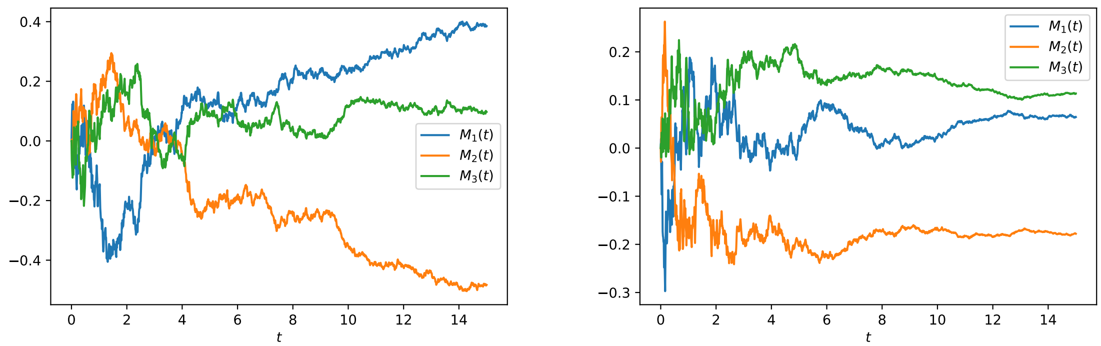
Alternatively, the A-dimensional generator $\tilde{L}_s$ can be rewritten as

$$\tilde{L}_s f(x) = \sum_{\substack{i,j=1\\ i<j}}^{A} \frac{p_i(Z(s))p_j(Z(s))}{2(1+s)^2}\left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j}\right)^2 f(x), \quad x \in \Delta_{A-1},$$

where $\left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j}\right)^2 := \frac{\partial^2}{(\partial x_i)^2} + \frac{\partial^2}{(\partial x_j)^2} - 2\frac{\partial^2}{\partial x_i \partial x_j}$ is the second derivative along the diagonal $x_i = x_j$. From this form of the generator it is easy to see (e.g. by a coordinate transformation) that M solves the system of stochastic differential equations

$$dM_i(t) = \sum_{j\neq i} \frac{\sqrt{p_i(Z(t))p_j(Z(t))}}{1+t} dB_{i,j}(t)\ , \quad i = 1,\ldots,A \tag{4.14}$$

where $B_{i,j}$ is a standard Brownian motion, which is independent of $B_{k,l}$ if $\{i,j\} \neq \{k,l\}$ and $B_{j,i} = -B_{i,j}$ for $i \neq j$. It follows immediately that $\left(\sum_{i=1}^{A} dM_i(t)\right) = 0$ for all $t > 0$.



(a) $F_1(k) = F_2(k) = F_3(k) = \sqrt{k}$, $\chi(0) = \left(\frac{8}{10}, \frac{1}{10}, \frac{1}{10}\right)$, Here $\lim_{t\to\infty}\langle M_1\rangle_t \approx 0.2474$.

(b) $F_1(k) = F_2(k) = F_3(k) = k^2$, $\chi(0) = \left(\frac{2}{5}, \frac{3}{10}, \frac{3}{10}\right)$, Here $\lim_{t\to\infty}\langle M_1\rangle_t \approx 0.1908$.

Figure 4.3: Realisations of the process M for different feedback functions and $A = 3$ generated by the Euler-Maruyama method for (4.14) with bandwidth $\frac{1}{100}$.

Hence, the sum $\sum_{i=1}^{A} M_i(t) = 0$ is a conserved quantity. Consequently, the state space of M is the tagent space $T\Delta_{A-1}$. This allows the following interpretation of the limit process M: Each pair of agents exchanges mass according to a time-changed Brownian motion and the exchange of several distinct pairs of agents is independent. Figure 4.3 finally shows two simulations of the process M with polynomial feedback.

4.2.2 Convergence of the predictable part

In order to complete the proof of Theorem 4.4, let us now turn to the predictable part $H^{(N)}$ in the Doob decomposition (4.4), which accounts for the drift part of (4.10). It is important to recall that $H^{(N)}(\lfloor Nt\rfloor)$ is deterministic when $M^{(N)}(\lfloor Ns\rfloor)$ is given for $s \le t$. Because of that, it is possible to express the limit process of $\sqrt{N}\big(\chi(0) + H^{(N)}(\lfloor Nt\rfloor) - Z(t)\big)$ for $N \to \infty$ in terms of the limit M of $\sqrt{N}M^{(N)}$. In Section 4.1, we derived that $\chi(0) + H^{(N)}$ converges to the deterministic process Z for $N \to \infty$ and the following result describes the deviation under appropriate scaling.

Theorem 4.12. *Suppose that the assumptions of Theorem 4.4 are fulfilled. Then*

$$\sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt\rfloor) - Z(t)\right)_{t\ge 0} \xrightarrow{N\to\infty} (H(t))_{t\ge 0} \quad \text{weakly on } \mathbb{D}([0,\infty), T\Delta_{A-1}) \,,$$

where H is the solution of the system of random ordinary differential equations (RODE)

$$\frac{d}{dt}H(t) = \frac{DG(Z(t))}{1+t} \cdot (H(t) + M(t)) \,, \quad H(0) = 0 \,. \tag{4.15}$$

Here, $DG(z)\colon T\Delta_{A-1} \to \mathbb{R}^A$ denotes the differential operator of G at the point $z \in \Delta^o_{A-1} \subset \mathbb{R}^A$, i.e. $DG(z)\cdot x$ is the derivative of G at z in direction $x \in T\Delta_{A-1}$. Note that $H(t)$

as well as $M(t)$ (as described in the previous section) are in the tangent space $T\Delta_{A-1} \subset \mathbb{R}^A$ (4.8), and therefore also $H(t)+M(t) \in T\Delta_{A-1}$. If G is well defined on an open neighbourhood of Δ_{A-1} in $\mathbb{R}^A$ (like in Example 4.6), then DG can be interpreted as the common differential matrix and $\cdot$ as the matrix-vector product.

The solution of a RODE is defined pathwise, in the sense that for any fixed realisation $\omega \in \Omega$ $M(t) = M(t,\omega)$ is a deterministic function, such that $H(t) = H(t,\omega)$ is the solution of the ODE (4.15). Further details on the theory of RODEs can be found e.g. in [71].

Consequently for fixed $\omega \in \Omega$, (4.15) is a linear, time-inhomogeneous ODE, the solution of which can be expressed as the matrix exponential

$$H(t) = e^{\int_0^t \frac{DG(Z(s))}{1+s}ds} \int_0^t e^{-\int_0^s \frac{DG(Z(u))}{1+u}du} \frac{DG(Z(s))}{1+s} \cdot M(s)ds.$$

An important part of the proof of Theorem 4.12 will be the tightness of the sequence of processes $\sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt \rfloor) - Z(t)\right)_{t\geq 0}$. For that, we bound its increments by the supremum of the martingale $M^{(N)}$.

Lemma 4.13. *In the situation of Theorem 4.12 we have with probability one for all $0 \leq s < t \leq T$*

$$\begin{aligned}\|H^{(N)}(\lfloor Nt \rfloor)-&Z(t) - H^{(N)}(\lfloor Ns \rfloor) + Z(s)\| \\ &\leq const. \left((t-s) \sup_{0\leq u\leq T} \|M^{(N)}(\lfloor Nu \rfloor)\| + \frac{t-s}{\sqrt{N}} + \frac{1}{N}\right),\end{aligned}$$

where const. is a constant only depending on G and T.

Proof. Let $L > 0$ be a Lipschitz constant for G (at least on a sufficiently large compact subset of Δ^o_{A-1}). We use (4.4) and calculate:

$$\begin{aligned}
&\|H^{(N)}(\lfloor Nt \rfloor) - Z(t) - H^{(N)}(\lfloor Ns \rfloor) - Z(s)\| \\
&= \left\| \sum_{k=\lfloor Ns \rfloor}^{\lfloor Nt \rfloor -1} \frac{G(N+k, \chi^{(N)}(k))}{N+k+1} - \int_s^t \frac{G(Z(u))}{1+u}du \right\| \\
&\leq \left\| \sum_{k=\lfloor Ns \rfloor}^{\lfloor Nt \rfloor -1} \frac{G(\chi^{(N)}(k))}{N+k+1} - \int_s^t \frac{G(Z(u))}{1+u}du \right\| \\
&\quad + \left\| \sum_{k=\lfloor Ns \rfloor}^{\lfloor Nt \rfloor -1} \frac{G(N+k, \chi^{(N)}(k))}{N+k+1} - \sum_{k=\lfloor Ns \rfloor}^{\lfloor Nt \rfloor -1} \frac{G(\chi^{(N)}(k))}{N+k+1} \right\| \\
&\leq \left\| \int_s^t \frac{G(\chi^{(N)}(\lfloor Nu \rfloor)) - G(Z(u))}{1+u} du \right\| + \frac{const.}{N} + \frac{const.}{\sqrt{N}} \sum_{k=\lfloor Ns \rfloor}^{\lfloor Nt \rfloor -1} \frac{1}{N+k+1} \\
&\leq \int_s^t \frac{\|G(\chi^{(N)}(\lfloor Nu \rfloor)) - G(Z(u))\|}{1+u} du + \frac{const.}{N} + const. \frac{t-s}{\sqrt{N}}
\end{aligned}$$

$$
\begin{aligned}
&\leq L \int_s^t \frac{\|\chi^{(N)}(\lfloor Nu \rfloor) - Z(u)\|}{1+u} du + const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\leq L \int_s^t \|\chi^{(N)}(\lfloor Nu \rfloor) - Z(u)\| du + const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\leq L \int_s^t \|\chi(0) + H^{(N)}(\lfloor Nu \rfloor) - Z(u)\| du + L \int_s^t \|M^{(N)}(\lfloor Nu \rfloor)\| du + const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\leq L \int_s^t \|H^{(N)}(\lfloor Nu \rfloor) - Z(u) - H^{(N)}(\lfloor Ns \rfloor) + Z(s)\| du + const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\quad + L \int_s^t \|\chi(0) + H^{(N)}(\lfloor Ns \rfloor) - Z(s)\| du + L(t-s) \sup_{0 \leq u \leq T} \|M^{(N)}(\lfloor Nu \rfloor)\| \\
&= L \int_s^t \|H^{(N)}(\lfloor Nu \rfloor) - Z(u) - H^{(N)}(\lfloor Ns \rfloor) + Z(s)\| du + const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\quad + L(t-s) \|\chi(0) + H^{(N)}(\lfloor Ns \rfloor) - Z(s)\| + L(t-s) \sup_{0 \leq u \leq T} \|M^{(N)}(\lfloor Nu \rfloor)\|
\end{aligned}
$$

In line 2, the second summand is of order $1/\sqrt{N}$ due to assumption (4.9). Now Grönwall's inequality yields:

$$
\begin{aligned}
\|H^{(N)}(\lfloor Nt \rfloor) - Z(t) - H^{(N)}(\lfloor Ns \rfloor) + Z(s)\| &\leq e^{L(t-s)} \cdot \Big(const. \frac{t-s}{\sqrt{N}} + \frac{const.}{N} \\
&\quad + L(t-s) \|\chi(0) + H^{(N)}(\lfloor Ns \rfloor) - Z(s)\| + L(t-s) \sup_{0 \leq u \leq T} \|M^{(N)}(\lfloor Nu \rfloor)\| \Big)
\end{aligned}
$$

Repeating the same calculation with 0 in the place of s and s instead of t yields:

$$
\|\chi(0) + H^{(N)}(\lfloor Ns \rfloor) - Z(s)\| \leq e^{Ls} \cdot \left(Ls \sup_{0 \leq u \leq T} \|M^{(N)}(\lfloor Nu \rfloor)\| + const. \frac{s}{\sqrt{N}} + \frac{const.}{N} \right)
$$

Combining these two inequalities proves the claim. □

We are now ready for the proof of Theorem 4.12.

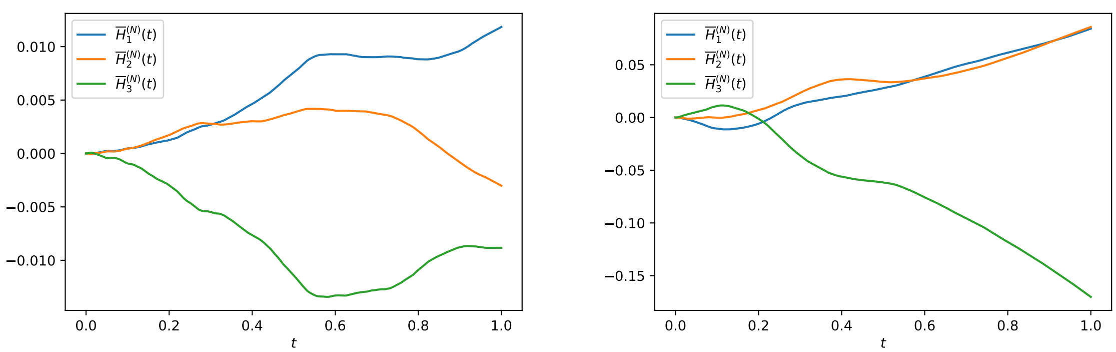
Proof. Via [77, Proposition VI.3.26], we get tightness of $\left(\sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt \rfloor) - Z(t)\right)_{t \geq 0}\right)_N$ from Lemma 4.13 and the stochastic boundedness of the sequence $(\sqrt{N} M^{(N)}(\lfloor Nt \rfloor))_{t \geq 0}$ (see proof of Lemma 4.8). Now we show that the limit of any convergent subsequence is as desired. For simplicity of notation, assume that the sequence is convergent itself. Since Theorem 4.7 applies, we can take an appropriate probability space Ω, such that the convergence $\sqrt{N} M^{(N)}(\lfloor Nt \rfloor, \omega) \xrightarrow{N \to \infty} M(t, \omega)$ holds locally uniformly almost surely. Note that this already implies $Z^{(N)}(\omega) \xrightarrow{N \to \infty} Z$ locally uniformly. Now, fix

$\omega \in \Omega$. Using (4.4) and the mean value theorem, we get

$$\begin{aligned}
&\sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt \rfloor) - Z(t)\right) = \sqrt{N}\left(\sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{G(N+k, \chi^{(N)}(k))}{N+k+1} - \int_0^t \frac{G(Z(s))}{1+s} ds\right) \\
&= \sqrt{N}\left(\sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{G(\chi^{(N)}(k))}{N+k+1} - \int_0^t \frac{G(Z(s))}{1+s} ds + \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{G(N+k, \chi^{(N)}(k)) - G(\chi^{(N)}(k))}{N+k+1}\right) \\
&= \sqrt{N}\left(\int_0^t \frac{G(\chi^{(N)}(\lfloor Ns \rfloor)) - G(Z(s))}{1+s} ds + O\left(\frac{1}{N}\right) + o\left(\frac{1}{\sqrt{N}}\right) \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N+k+1}\right) \\
&= \sqrt{N} \int_0^t \frac{G(\chi^{(N)}(\lfloor Ns \rfloor)) - G(Z(s))}{1+s} ds + o(1) \\
&= \sqrt{N} \int_0^t \frac{DG(m^{(N)}(s)) \cdot (\chi^{(N)}(\lfloor Ns \rfloor) - Z(s))}{1+s} ds + o(1) \\
&= \int_0^t DG(m^{(N)}(s)) \cdot \frac{\sqrt{N}\left((\chi(0) + H^{(N)}(\lfloor Ns \rfloor) - Z(s) + M^{(N)}(\lfloor Ns \rfloor)\right)}{1+s} ds + o(1) \\
&\xrightarrow{N \to \infty} \int_0^t \frac{DG(Z(s)) \cdot (H(s) + M(s))}{1+s} ds,
\end{aligned}$$

where $m^{(N)}(s)$ is an intermediate value between $Z(s)$ and $\chi^{(N)}(\lfloor Ns \rfloor)$. In line 3, we used assumption (4.9) once again. The claim follows since (4.15) has a unique solution due to the Theorem of Picard-Lindelöf, and $H(t) \in T\Delta_{A-1}$ since $H^{(N)}(k) \in T\Delta_{A-1}$ for all $N \geq 1$. □

Figure 4.4 shows a simulation of the process $\sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt \rfloor) - Z(t)\right)$ for large N and small t. Note that the limit process (4.15) has continuously differentiable paths, their regularity is equivalent to that of an integrated Brownian motion. As a consequence of Proposition 4.5, $H(t)$ is convergent for $t \to \infty$ with random limit $-\lim_{t\to\infty} M(t)$ in generic examples. Combining Theorem 4.11 and Theorem 4.12 yields the desired CLT for the difference $Z^{(N)} - Z = \chi(0) + H^{(N)} - Z + M^{(N)}$.



(a) $F_1(k) = F_2(k) = F_3(k) = \sqrt{k}$, $\chi(0) = \left(\frac{1}{10}, \frac{1}{10}, \frac{4}{5}\right)$

(b) $F_1(k) = F_2(k) = F_3(k) = k^2$, $\chi(0) = \left(\frac{3}{10}, \frac{3}{10}, \frac{2}{5}\right)$

Figure 4.4: The processes $\bar{H}^{(N)}(t) = (\bar{H}_1^{(N)}(t), \bar{H}_2^{(N)}(t), \bar{H}_3^{(N)}(t)) := \sqrt{N}\left(\chi(0) + H^{(N)}(\lfloor Nt \rfloor) - Z(t)\right)$ for $A = 3$ and $N = 100.000$.

Chapter 5

The Wealth of Losers

In this chapter, we assume that at least one agent satisfies (M), such the process exhibits strong monopoly (Theorem 2.2). In Chapter 3 we presented results on the prediction of the monopolist. Now, we turn to the losers and study the tail distribution of the number of steps won by a losing agent. To do that, we first derive a general result on the tail behaviour of birth processes at a fixed observation time conditioned on non-explosion (Section 5.1). In Section 5.2, we apply this finding to extend results in [132, 102, 44] (in particular Theorem 2.8) on the wealth of losers to situations with more than two agents with possibly different feedback functions. Moreover, we characterize the tail dependence of several losing agents and discuss applications, which encompass e.g. the total wealth of all losers or the time to monopoly. This also allows for an analysis of the large deviation behaviour of the process. What may seem paradoxical at first is that losers with feedback close to the transition (M) are most likely to win in many steps. In order to provide a detailed description of this transition, we finally also consider the wealth of losers which do not satisfy (M) (Section 5.3). We also discuss some special features of large systems with diverging size $A \to \infty$, but we refer to Appendix D for details.

For later convenience, we establish the following notation for real sequences (x_k) and (y_k):

$$
\begin{aligned}
x_k \sim y_k \quad &\Leftrightarrow \quad \lim_{k\to\infty} \frac{x_k}{y_k} = 1\,, \\
x_k \prec y_k \quad &\Leftrightarrow \quad \limsup_{k\to\infty} \frac{x_k}{y_k} < \infty\,, \\
x_k \asymp y_k \quad &\Leftrightarrow \quad x_k \prec y_k \text{ and } y_k \prec x_k\,.
\end{aligned}
\tag{5.1}
$$

The results of this chapter have been published in [68].

5.1 Tails of explosive birth processes

Birth processes are fundamental models for growth processes, having applications in numerous fields such as biology [47], chemistry [58], computer science [12] or reliability theory [1]. The simple birth process was first introduced by Yule [131] in 1924 in connection with evolutionary theory and independently in 1937 by Furry [64] in a physical context, which is why it is also known as Yule-Furry process. Birth and death processes, which were introduced by Feller [55] in 1939, pose an important and well-studied extension of this model.

Given a rate function $F\colon \mathbb{N} := \{1, 2, \ldots\} \to (0, \infty)$, a **(pure) birth process** is a homogeneous, continuous-time Markov process $(\Xi(t))_{t\geq 0}$ on the state space $\mathbb{N}$, which jumps from k to $k+1$ at rate $F(k)$. More formally, take independent random variables $\tau(k)$, which are exponentially distributed with parameter $F(k)$. Then a birth process with initial condition $\Xi(0) \in \mathbb{N}$ is the corresponding counting process with sojourn times $\tau(k)$, i.e.

$$\Xi(t) := \min\left\{ k \in \mathbb{N}\colon \sum_{l=\Xi(0)}^{k} \tau(l) > t \right\} \quad \text{for all } t \in [0, \infty)\,, \tag{5.2}$$

where $\min \emptyset := \infty$. A widely studied case is the simple (or linear) birth process, where $F(k) = \lambda k$ for some $\lambda > 0$. $F(k) = \lambda$ corresponds to a homogeneous Poisson process.

Since F is strictly positive we have $\Xi(t) \to \infty$ as $t \to \infty$ almost surely. Ξ is called **explosive** if $\Xi(t) = \infty$ for some finite $t < \infty$, which is obviously equivalent to $\sum_{l=\Xi(0)}^{\infty} \tau(k) < \infty$ and we define the explosion time

$$T := \sum_{k=\Xi(0)}^{\infty} \tau(k) < \infty \quad \text{with density } g \tag{5.3}$$

in the explosive case. The existence and properties of g will be discussed in Lemma 5.2 below. The well-established Theorem of Feller-Lundberg [56, 91] states that

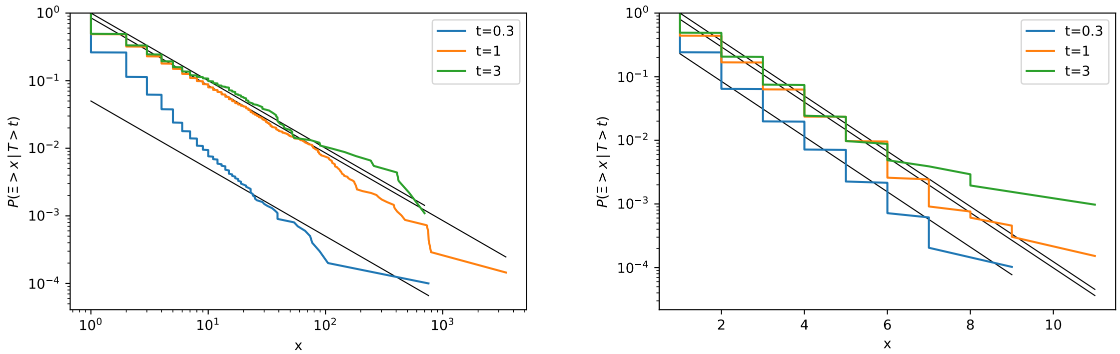
$$\Xi \text{ is explosive (i.e. } T < \infty) \text{ if and only if } \quad \sum_{k=1}^{\infty} \frac{1}{F(k)} < \infty\,,$$

i.e. (M) holds. [56] further contains an explicit formula for the distribution of $\Xi(t)$ for fixed t,

$$\mathbb{P}(\Xi(t) = x) = \sum_{k=\Xi(0)}^{x} \frac{\prod_{l=\Xi(0)}^{x-1} F(l)}{\prod_{\substack{l=\Xi(0)\\ l\neq k}}^{x} (F(l) - F(k))} e^{-F(k)t} \quad \text{for } x \in \mathbb{N}\,. \tag{5.4}$$

Remarkably, this holds for both explosive and non-explosive birth processes, and for the non-explosive case. [104] provides a formula for the expectation of $\Xi(t)$. For linear birth processes with $F(k) = k$, $\Xi(t)$ has a negative binomial distribution and grows exponentially in t (see e.g. [13, 89]). Another result worth mentioning in the context of explosive birth processes is [105, Theorem 3.1], which determines the asymptotics of $\Xi(T - t)$ for $t \to 0$. The asymptotics for $t \to \infty$ of non-explosive birth processes are described in e.g. [127, 126, 11].

This section is dedicated to the explosive case, where we characterize the tail of $\Xi(t)$ for fixed $t > 0$ conditioned on not having exploded yet, i.e. $T > t$. This question was addressed in a recent paper [59] based on (5.4) and focusing on numerical studies. We choose a different approach to get a rigorous and fairly general result (Theorem 5.1).



(a) $F(k) = k^2$ (b) $F(k) = e^{k-1}$

Figure 5.1: The empirical distribution of $\Xi(t)$ conditioned on $T > t$ for different F and t. The black lines show the predicted tail according to Theorem 5.1. 10,000 realizations of the process (5.2) were simulated each. In (a) 10 (resp. 3112, resp. 9088) have already exploded at time $t = 0.3$ (resp. $t = 1$, resp. $t = 3$) and in (b) 30 (resp. 3424, resp. 8974). As the criterion for explosion, the simulation was stopped after 10^6 summands in (a) and 100 summands in (b).

5.1.1 Main result

Using the notation established above, we directly formulate our main result on the tail of birth processes conditioned on non-explosion.

Theorem 5.1. *Assume that (M) is fulfilled. Then*

$$\mathbb{P}(\Xi(t) = x) \sim \frac{g(t)}{F(x)} \quad \textit{for } x \to \infty\,,$$

holds for all $t > 0$, where g is the density of the explosion time T (5.3). For any $t_0 > 0$, the convergence is uniform in $t > t_0$. Consequently, we get the following conditional tail distribution:

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \sim -\frac{d}{dt}\log(\mathbb{P}(T > t)) \sum_{k=x+1}^{\infty} \frac{1}{F(k)} \quad \textit{for } x \to \infty$$

Figure 5.1 numerically illustrates Theorem 5.1 for quadratic and exponential rate functions. Indeed, when t is moderate or large, then the approximation seems good even for small x. The approximation is rather inaccurate for small t, when $\Xi(t)$ is likely to be close to $\Xi(0)$. Theorem 5.1 separates the tail of $\Xi(t)$ into a prefactor that depends only on t and a sequence that depends only on x. The prefactors $-\frac{d}{dt}\log\mathbb{P}(T > t) = \frac{g(t)}{\mathbb{P}(T>t)}$ are plotted in Figure 5.2 for several feedback functions and will be discussed in the following.

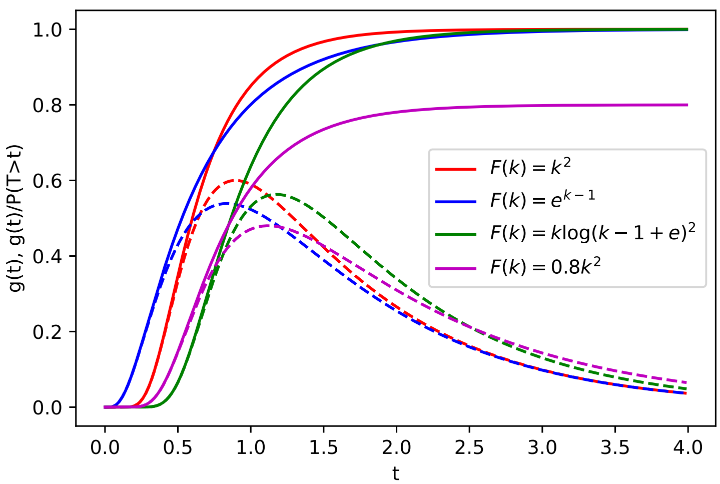


Figure 5.2: The factors $-\frac{d}{dt}\log\mathbb{P}(T > t) = \frac{g(t)}{\mathbb{P}(T>t)}$ (full lines) and $g(t)$ (dotted lines) from Theorem 5.1 for different feedback functions. The full lines converge to $F(1)$ for $t \to \infty$, as discussed later in (5.13). For the numerical computation, g was approximated by g_n from Lemma 5.3 with $n = 100$.

In the proof of Theorem 5.1, we will be particularly interested in first passage times, which we formally define as

$$T(k) \coloneqq \sum_{l=\Xi(0)}^{k} \tau(l) \quad \text{with density } g_k$$

and the remaining time until explosion

$$\bar{T}(k) \coloneqq \sum_{l=k+1}^{\infty} \tau(l) \quad \text{with density } \bar{g}_k$$

for $k \geq \Xi(0)$. We start the proof with an auxiliary lemma on analytic properties of the density g.

Lemma 5.2. *If (M) is satisfied, then the density g of the explosion time T (5.3) exists and has the following properties.*

1. *g is bounded with $g(0) = 0$ and $g(t) > 0$ for all $t > 0$.*
2. *g has a unique maximum at $m > 0$, such that $g(t)$ is increasing for $t < m$ and decreasing for $t > m$.*
3. *For any $t_0 > 0$, the logarithmic derivative $\left|\frac{d}{dt}\log g(t)\right|$ is bounded uniformly in $t \geq t_0$.*
4. *g is not analytic in 0. In particular, $\frac{d^k}{dt^k}g(0) = 0$ holds for all $k \geq 1$*

Proof. For simplicity of notation, assume $\Xi(0) = 1$. Existence of g follows e.g. from the inversion formula for characteristic functions. Since $g_2(0) = 0$ and $\mathbb{P}(T(2) \leq x) \geq \mathbb{P}(T \leq x)$, we have $g(0) = 0$. From the convolution formula for $T(1) + \bar{T}(1)$, i.e.

$$g(t) = \int_0^t g_1(t-s)\bar{g}_1(s)\,ds = F(1)\int_0^t e^{-F(1)(t-s)}\bar{g}_1(s)\,ds\,, \tag{5.5}$$

we can conclude that g is bounded by $\max_{t\geq 0} g_1(t) = F(1)$. From Markov's inequality, we get that for any $\epsilon > 0$ and k large enough

$$\mathbb{P}(\bar{T}(k) < \epsilon/2) > 0\,,$$

such that $\mathbb{P}(T < \epsilon) > 0$ holds, too. Hence, $g(t) > 0$ for small $t > 0$ and analogously $\bar{g}_1(t) > 0$ for small $t > 0$. Then $g(t) > 0$ for any $t > 0$ follows from the convolution formula (5.5).

For 2., we show that for all $k \in \mathbb{N}$ there is $m_k \geq 0$ such that $g_k'(t) > 0$ for all $t < m_k$ and $g_k'(t) < 0$ for all $t > m_k$. This implies 2. since $g_k(t) \to g(t)$ for $k \to \infty$ uniformly in $t \geq 0$ (see [26] for a rigorous argument). We show this by induction over k. Assume that g_k has the desired properties. Then we get from the convolution formula for $T(k+1) = T(k) + \tau(k+1)$

$$g_{k+1}'(t) = \int_0^t F(k+1)e^{-F(k+1)s}g_k'(t-s)ds + F(k+1)e^{-F(k+1)t}g_k'(0)\,. \tag{5.6}$$

For small enough $t > 0$, it follows $g_{k+1}'(t) > 0$ since $g_k'(t) > 0$. Define

$$m_{k+1} \coloneqq \min\{t > 0\colon g_{k+1}'(t) = 0\} \tag{5.7}$$

and rephrase

$$\begin{aligned}
\frac{g_{k+1}'(t)}{F(k+1)} &= \int_0^{t-m_{k+1}} e^{-F(k+1)s}g_k'(t-s)ds + \int_{t-m_{k+1}}^t e^{-F(k+1)s}g_k'(t-s)ds + e^{-F(k+1)t}g_k'(0)\\
&= \int_0^{t-m_{k+1}} e^{-F(k+1)s}g_k'(t-s)ds\\
&\quad + e^{-F(k+1)(t-m_{k+1})}\left(\int_0^{m_{k+1}} e^{-F(k+1)s}g_k'(m_{k+1}-s)ds + e^{-F(k+1)m_{k+1}}g_k'(0)\right)\\
&= \int_0^{t-m_{k+1}} e^{-F(k+1)s}g_k'(t-s)ds
\end{aligned}$$

for $t > m_{k+1}$, using (5.6) and (5.7). Since necessarily $g_k'(t) < 0$ for $t > m_{k+1} > m_k$, we get $g_{k+1}'(t) < 0$ for $t > m_{k+1}$.

For 3., it only remains to show $\liminf_{t\to\infty} \frac{d}{dt}\log g(t) > -\infty$ together with 1. and 2.. This follows from

$$\frac{d}{dt}\log g(t) = \frac{g'(t)}{g(t)} = \frac{-F(1)^2\int_0^t e^{-F(1)(t-s)}\bar{g}_1(s)ds + F(1)\bar{g}_1(t)}{\int_0^t F(1)e^{-F(1)(t-s)}\bar{g}_1(s)ds} > -F(1)\,.$$

For 4., we first show that $\frac{d^k}{dt^k} g_{k+2}(0) = 0$ via induction over $k \geq 1$. Assume that $\frac{d^{k-1}}{dt^{k-1}} g_{k+1}(0) = 0$, which is easy to check for $k = 1$. Then:

$$\begin{aligned}
\frac{d^k}{dt^k} g_{k+2}(t) &= F(k+1) \frac{d}{dt} \int_0^t e^{-F(k+1)s} \frac{d^{(k-1)}}{dt^{(k-1)}} g_{k+1}(t-s) ds \\
&= F(k+1) \frac{d}{dt} \int_0^t e^{-F(k+1)(t-s)} \frac{d^{(k-1)}}{dt^{(k-1)}} g_{k+1}(s) ds \\
&= -F(k+1)^2 \int_0^t e^{-F(k+1)(t-s)} \frac{d^{(k-1)}}{dt^{(k-1)}} g_{k+1}(s) ds + F(k+1) \frac{d^{(k-1)}}{dt^{(k-1)}} g_{k+1}(t) \xrightarrow{t \to 0} 0
\end{aligned}$$

Now, assume that there is $k \geq 1$ such that $\frac{d^k}{dt^k} g(0) > 0$. Then $g_{k+2}(t) < g(t)$ for small enough $t > 0$ since $\frac{d^k}{dt^k} g_{k+2}(0) = 0$. On the other hand, $g_{k+2}(t) > g(t)$ for small enough $t > 0$ due to $\mathbb{P}(T(k+2) \leq t) > \mathbb{P}(T \leq t)$. Hence, $\frac{d^k}{dt^k} g(0) = 0$. □

For injective $F \colon \mathbb{N} \to (0, \infty)$, [132] provides an explicit formula for $g(t)$, which is useful for numerical approximation (see Figure 5.2).

Lemma 5.3. *[132, Lemma 3.2.1., Lemma 3.2.2.] Let $F(1), \dots, F(n)$ be positive and pairwise distinct with $\tau(k)$ exponentially distributed and independent with parameters $F(k)$. Then $T(n) = \sum_{k=1}^n \tau(k)$ has the density function*

$$g_n(x) = \left(\prod_{k=1}^n F(k) \right) \sum_{k=1}^n \frac{e^{-F(k)x}}{\prod_{l \neq k} (F(l) - F(k))} \quad \text{for } x \geq 0 \, .$$

The formula also applies for $n = \infty$ and $g(x)$ if (M) *holds, i.e. $\sum_{k=1}^\infty \frac{1}{F(k)} < \infty$.*

We continue our proof of Theorem 5.1 with another technical lemma.

Lemma 5.4. *Assume that (M) is fulfilled and that F is strictly monotone. Then we have for any $t_0 > 0$ that*

$$\frac{g_k(t)}{g(t)} \to 1 \quad \text{for } k \to \infty \text{ uniformly in } t \geq t_0 \, .$$

Proof. Let $\epsilon > 0$. For all $t_0 > 0$, there is $\delta > 0$ such that $|g(t+s)/g(t) - 1| < \epsilon$ for $0 < s < \delta$ and all $t > t_0$, because the logarithmic derivative of g is bounded (see Lemma 5.2). Let k be large enough such that $\mathbb{P}(\bar{T}(k) > \delta) < \epsilon$. Then we have by the convolution formula for $T(k) = T - \bar{T}(k)$

$$\begin{aligned}
\left| \frac{g_k(t)}{g(t)} - 1 \right| &= \left| \frac{1}{g(t)} \int_t^\infty g(s) \bar{g}_k(s-t) ds - 1 \right| \\
&\leq \left| \frac{1}{g(t)} \int_t^{t+\delta} g(s) \bar{g}_k(s-t) ds - 1 \right| + \left| \frac{1}{g(t)} \int_{t+\delta}^\infty g(s) \bar{g}_k(s-t) ds \right| \\
&\leq 2\epsilon + \epsilon \frac{\max_{s \geq t+\delta} g(s)}{g(t)}
\end{aligned}$$

for all $t > 0$. Properties 1. and 2. of Lemma 5.2 complete the proof. □

Note that obviously g_k converges to g globally uniformly for $t \geq 0$, but Lemma 5.4 is a stronger assertion. We are now ready to finish the proof of Theorem 5.1.

Proof of Theorem 5.1. By the independence of $T(k)$ and $\tau(k+1)$, we get:

$$\begin{aligned}\mathbb{P}(\Xi(t) = x) &= \mathbb{P}(T(x-1) < t,\, T(x) \geq t) = \mathbb{P}(t - \tau(x) \leq T(x-1) < t)\\ &= F(x) \int_0^\infty \mathbb{P}(t - u \leq T(x-1) < t)\, e^{-F(x)u}\, du\\ &= F(x) \int_0^{a_x} \mathbb{P}(t - u \leq T(x-1) < t)\, e^{-F(x)u}\, du\\ &\quad + F(x) \int_{a_x}^\infty \mathbb{P}(t - u \leq T(x-1) < t)\, e^{-F(x)u}\, du\,,\end{aligned}$$

where $a_x := \frac{1}{\sqrt{F(x)}}$. Using the mean value theorem and the uniform asymptotic $g_x \sim g$ for $x \to \infty$ (see Lemma 5.4), this implies for some intermediate value $\tilde{t}(u) \in (t-u, t)$

$$\mathbb{P}(t - u \leq T(x) < t) = g_x(\tilde{t}(u))u \sim g(t)u \quad \text{for } u \to 0,\, x \to \infty\,. \tag{5.8}$$

Hence, the first integral is asymptotically

$$\begin{aligned}F(x) \int_0^{a_x} \mathbb{P}(t - u \leq T(x-1) < t)\, e^{-F(x)u}\, du &\sim g(t) \int_0^{a_x} u\, F(x) e^{-F(x)u}\, du\\ = -g(t)\left(a_x + \frac{1}{F(x)}\right) e^{-F(x)a_x} + \frac{g(t)}{F(x)} &\sim \frac{g(t)}{F(x)} \quad \text{for } x \to \infty\,.\end{aligned}$$

The second integral is bounded by

$$F(x) \int_{a_x}^\infty \mathbb{P}(t - u \leq T(x-1) < t)\, e^{-F(x)u}\, du \leq \int_{a_x}^\infty F(x)\, e^{-F(x)u}\, du = e^{-F(x)a_x}\,,$$

which converges to zero faster than $\frac{1}{F(x)}$ for $x \to \infty$. □

5.1.2 Discussion and extensions

According to Theorem 5.1, the tail of $\Xi(t)$ conditioned on non-explosion $T > t$ is heavier the slower the feedback function F increases. This may seem paradoxical at first and can be explained as follows: For fast increasing F, if conditioned on T to be large, it is likely that the first summands of T are large and the chain spends a long time in states with the smallest exit rates so that $\Xi(t)$ is rather small. This corresponds to the general principle in large deviation theory that the rare event $T > t$ (for $t \to \infty$) is realized in the most likely (or least unlikely) way. We get back to this in more detail in Section 5.1.3. On the other hand, the prefactor $-\frac{d}{dt}\log(\mathbb{P}(T > t))$ for small t in Theorem 5.1 is significantly smaller when F increases slowly, as is visible in Figure 5.2.

Theorem 5.1 also directly implies

$$\mathbb{E}[\Xi(t)^r \,|\, T > t] < \infty \quad \Leftrightarrow \quad \sum_{k=1}^\infty \frac{k^r}{F(k)} < \infty \tag{5.9}$$

for any $r > 0$. Let us now discuss some interesting examples.

Example 5.5. 1. Let $F(k) = k^\beta$ for $\beta > 1$. Then

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \sim \frac{-1}{\beta - 1} \frac{d}{dt} \log(\mathbb{P}(T > t))\, x^{1-\beta} \quad \text{for } x \to \infty \tag{5.10}$$

is a power-law distribution with exponent $1 - \beta$, and

$$\mathbb{E}[\Xi(t)^r \,|\, T > t] < \infty \quad \Leftrightarrow \quad r < \beta - 1\,.$$

2. Let $F(k) = e^{\beta k}$ for $\beta > 0$. Then

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \sim -\frac{e}{\beta(e-1)} \frac{d}{dt} \log(\mathbb{P}(T > t))\, e^{-\beta(x+1)} \quad \text{for } x \to \infty$$

has an exponential tail and $\mathbb{E}[\Xi(t)^r \,|\, T > t] < \infty$ for all $r > 0$. Here and in the first example the tail is lighter the larger β is.

3. Let $F(k) = k(\log k)^\beta$ for $\beta > 1$. Then

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \sim \frac{-1}{\beta - 1} \frac{d}{dt} \log(\mathbb{P}(T > t))\, (\log x)^{1-\beta} \quad \text{for } x \to \infty \tag{5.11}$$

decays logarithmically slowly and $\mathbb{E}[\Xi(t)^r \,|\, T > t] = \infty$ for all $r > 0$.

Obviously we have

$$\lim_{t\to 0} \mathbb{P}(\Xi(t) = \Xi(0) \,|\, T > t) = 1$$

for all feedback functions, but remarkably the tail asymptotics for $\Xi(t)$ hold for any positive $t > 0$, in particular for polynomially growing feedback it has a heavy-tailed distribution. This is consistent with vanishing prefactors in Theorem 5.1, i.e.

$$\lim_{t\to 0} \frac{d}{dt} \log \mathbb{P}(T > t) = 0 \quad \text{and} \quad g(0) = 0\,,$$

which is established in Lemma 5.2 and illustrated in Figure 5.2. Since, $\Xi(t) = \infty$ is likely for large t, the prefactor $g(t)$ of the mass function vanishes for large t, while that of the tail conditioned on non-explosion $T > t$ converges to the minimum of F (see (5.13) in Section 5.1.3).

By replacing the mean value theorem in (5.8) by second order Taylor expansion,

$$\mathbb{P}(t - u < T(x) < t) = g_x(t)u + \frac{1}{2} g'_x(\tilde{t}(u))u^2\,,$$

we can also determine the **rate of convergence** in Theorem 5.1:

$$\mathbb{P}(\Xi(t) = x) - \frac{g(t)}{F(x)} \sim \frac{g'(t)}{2} \int_0^\infty u^2\, F(x)\, e^{-F(x)u} du = \frac{g'(t)}{F(x)^2} \quad \text{for } x \to \infty$$

Furthermore, we can also derive an asymptotic upper bound, which is even globally uniform.

Corollary 5.6. *Assume that (M) is fulfilled. Then*

$$\mathbb{P}(\Xi(t) = x) \prec \frac{g_x(t)}{F(x)} \quad \textit{for } x \to \infty$$

and

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \prec \frac{g_x(t)}{\mathbb{P}(T > t)} \sum_{k=x+1}^{\infty} \frac{1}{F(k)} \quad \textit{for } x \to \infty$$

holds globally uniformly in $t \geq 0$.

Proof. Using Theorem 5.1, we only have to derive the uniform asymptotic bound for small $t \leq t_0$. According to Lemma 5.2, there is $t_0 > 0$ such that g_x and g is increasing on the interval $(0, t_0)$. Then the proof of Theorem 5.1 applies analogously for $t \leq t_0$ if we just replace (5.8) by the (non-asymptotic) inequality

$$\mathbb{P}(t - u < T(x) < t) = g_x(\tilde{t}(u))u \leq g_x(t)u\,.$$

□

As a consequence of 4. in Lemma 5.2, $\lim_{t \to 0} \frac{g_k(t)}{g(t)} = \infty$ holds for all $k \geq 1$ and, consequently, the convergence in Lemma 5.4 is not uniform in $t \geq 0$. For this reason, there is no uniform lower bound corresponding to Corollary 5.6. In particular, the assumption $t_0 > 0$ is necessary in Theorem 5.1. This is already hinted at by the blue line in Figure 5.1 (a).

Applying Corollary 5.6, it is possible to **randomize the observation time** t in Theorem 5.1, which turns out to be useful for the application in Section 5.2.

Corollary 5.7. *Assume that (M) is fulfilled. Let S denote a positive random variable, which is independent of the process Ξ. Then as $x \to \infty$*

$$\mathbb{P}(\Xi(S) = x) \sim \frac{\mathbb{E}g(S)}{F(x)} \quad \textit{and} \quad P(\Xi(S) > x | T > S) \sim \frac{\mathbb{E}g(S)}{\mathbb{P}(T > S)} \sum_{k=x+1}^{\infty} \frac{1}{F(k)}\,.$$

Proof. Denote by $\mathbb{P}_S$ the law of S. Take $\epsilon > 0$. Then:

$$\mathbb{P}(\Xi(S) = x) = \int_0^\infty \mathbb{P}(\Xi(s) = x)\, d\mathbb{P}_S(s) = \int_0^\epsilon \mathbb{P}(\Xi(s) = x)\, d\mathbb{P}_S(s) + \int_\epsilon^\infty \mathbb{P}(\Xi(s) = x)\, d\mathbb{P}_S(s)$$

Let us now look at the second integral and apply Theorem 5.1.

$$\int_\epsilon^\infty \mathbb{P}(\Xi(s) = x)\, d\mathbb{P}_S(s) \sim \int_\epsilon^\infty \frac{g(s)}{F(x)}\, d\mathbb{P}_S(s) \xrightarrow{\epsilon \to 0} \frac{\mathbb{E}g(S)}{F(x)}$$

The first integral is asymptotically negligible due to Corollary 5.6:

$$\int_0^\epsilon \mathbb{P}(\Xi(s) = x)\, d\mathbb{P}_S(s) \prec \int_0^\epsilon \frac{g_x(s)}{F(x)}\, d\mathbb{P}_S(s) \xrightarrow{\epsilon \to 0} 0$$

The second asymptotic follows via

$$P(\Xi(S) > x | T > S) = \frac{\mathbb{P}(\Xi(S) > x,\, T > S)}{\mathbb{P}(T > S)} = \frac{\sum_{k=x+1}^{\infty} \mathbb{P}(\Xi(S) = k)}{\mathbb{P}(T > S)} \sim \frac{\mathbb{E}g(S)}{\mathbb{P}(T > S)} \sum_{k=x+1}^{\infty} \frac{1}{F(k)}$$

for $x \to \infty$. Note that $\Xi(S) = k$ implies $T > S$. □

5.1.3 Behaviour for $t \to \infty$

Figure 5.2 also indicates that $-\frac{d}{dt}\log(\mathbb{P}(T > t))$ is increasing in t and converges for $t \to \infty$. The latter observation can be rigorously confirmed for strictly monotone F, since we get from Lemma 5.3 that

$$g(t) \sim F(\Xi(0)) \left(\prod_{l=\Xi(0)+1}^{\infty} \frac{F(l)}{F(l) - F(\Xi(0))} \right) e^{-F(\Xi(0))t} \quad \text{for } t \to \infty$$

and by integration

$$\mathbb{P}(T > t) \sim \left(\prod_{l=\Xi(0)+1}^{\infty} \frac{F(l)}{F(l) - F(\Xi(0))} \right) e^{-F(\Xi(0))t} \quad \text{for } t \to \infty\,. \tag{5.12}$$

Thus,

$$\lim_{t\to\infty} -\frac{d}{dt}\log(\mathbb{P}(T > t)) = \lim_{t\to\infty} \frac{g(t)}{\mathbb{P}(T > t)} = F(\Xi(0))\,. \tag{5.13}$$

As a consequence, it stands to reason that $\lim_{t\to\infty} \mathbb{P}(\Xi(t) \in \cdot \,|\, T > t)$ defines a non-degenerated distribution. The following Proposition confirms this conjecture.

Proposition 5.8. *Assume that (M) is fulfilled and that F is strictly monotone. Then*

$$\mathbb{P}(\Xi(t) > x \,|\, T > t) \to 1 - \prod_{k=x+1}^{\infty} \left(1 - \frac{F(\Xi(0))}{F(k)}\right) \in (0,1) \quad \text{as } t \to \infty$$

for all $x \geq \Xi(0)$. Moreover, we have for any $s > 0$

$$\mathbb{P}(T - \tau(\Xi(0)) > s \,|\, T > t) \prec e^{-(F(\Xi(0)+1) - F(\Xi(0)))s} \quad \text{as } t \to \infty$$

uniformly in s.

Before we turn to the proof, we add some remarks.

1. The so-called **quasi-limiting distribution** $\nu^{\infty}_{\Xi(0)} := \lim_{t\to\infty} \mathbb{P}(\Xi(t) \in . \mid T > t)$ on $\{\Xi(0), \Xi(0)+1, \ldots\}$ has been studied in the context of Yaglom limits and quasi-stationary distributions for general Markov processes and in particular also for birth-death processes [36, 95]. In this limiting distribution, we can easily recover the tail $x \to \infty$ from Theorem 5.1

$$\lim_{t\to\infty} \mathbb{P}(\Xi(t) > x \,|\, T > t) = 1 - \prod_{k=x+1}^{\infty} \left(1 - \frac{F(\Xi(0))}{F(k)}\right) = 1 - \exp\left\{ \sum_{k=x+1}^{\infty} \log\left(1 - \frac{F(\Xi(0))}{F(k)}\right) \right\}$$
$$\sim 1 - \exp\left\{ -F(\Xi(0)) \sum_{k=x+1}^{\infty} \frac{1}{F(k)} \right\} \sim F(\Xi(0)) \sum_{k=x+1}^{\infty} \frac{1}{F(k)} \quad \text{as } x \to \infty\,.$$

Note that
$$\lim_{x\to\infty} 1 - \prod_{k=x+1}^{\infty}\left(1 - \frac{F(\Xi(0))}{F(k)}\right) = 0\,.$$
Hence, $\nu^{\infty}_{\Xi(0)}$ is a normalized probability distribution on $\{\Xi(0), \Xi(0)+1, \ldots\}$, in particular there is no probability mass in $x = \infty$. We give explicit examples for this distribution with power-law tails below.

2. If a birth-death process is irreducible except for one absorbing state (usually the state 0) with light-tailed hitting time, then quasi-limiting distributions are also quasi-stationary, i.e. do not change under the time evolution conditioned on non-absorption. Since we consider a pure birth process where all states are transient those results do not apply here. In particular, the quasi-limiting distributions $\nu^{\infty}_{\Xi(0)}$ depend on the initial condition and are not quasi-stationary. In fact, there are no quasi-stationary distributions for our process. Due to transience of all states, the only possible candidates would be $\delta_{\Xi(0)}$, but with Proposition 5.8 $\nu^{\infty}_{\Xi(0)}(\Xi(0)) < 1$ for all feedback functions.

3. But how does the event of non-explosion at some late time typically occur? According to Proposition 5.8, the time $\tau(\Xi(0))$ is of the same order as the explosion time T. Moreover, since the tail of T decays exponentially, $T - t$ has an exponential tail as well on the event $\{T > t\}$. Hence, conditioned on $\{T > t\}$ for large t, the first jump of the process happens close to t up to a light-tailed distribution (asymptotically independent of t), i.e.
$$\mathbb{P}(|\tau(\Xi(0)) - t| > s \,|\, T > t) \prec e^{-ms} \quad \text{as } t \to \infty$$
uniformly in s, where $m := \min\{F(\Xi(0)), F(\Xi(0)+1) - F(\Xi(0))\} > 0$. In particular this implies $\mathbb{P}(\Xi(\alpha t) = \Xi(0)|T > t) \to 1$ as $t \to \infty$ for all $\alpha < 1$. Nevertheless, there might be further jumps between the first one at $\tau(\Xi(0))$ and the observation time t with non-vanishing probability. Note that for $F(\Xi(0)+1) \approx F(\Xi(0))$ there can be several long sojourn times when $T > t$.

Proof of Proposition 5.8. Due to the strict monotonicity of F and Lemma 5.3, it is easy to see that
$$g_k(t) \sim F(\Xi(0)) \left(\prod_{l=\Xi(0)+1}^{k} \frac{F(l)}{F(l) - F(\Xi(0))}\right) e^{-F(\Xi(0))t} \quad \text{for } t \to \infty$$
for $k \geq \Xi(0)$ including $k = \infty$ with $g_\infty = g$. Then the first claim follows via de l'Hospital's Theorem:
$$\mathbb{P}(\Xi(t) > x \,|\, T > t) = 1 - \mathbb{P}(\Xi(t) \leq x \,|\, T > t) = 1 - \frac{\mathbb{P}(\Xi(t) \leq x)}{\mathbb{P}(T > t)} = 1 - \frac{\mathbb{P}(T(x) > t)}{\mathbb{P}(T > t)}$$
$$\sim 1 - \frac{g_x(t)}{g(t)} \xrightarrow{t\to\infty} 1 - \frac{F(\Xi(0))\left(\prod_{k=\Xi(0)+1}^{x} \frac{F(k)}{F(k)-F(\Xi(0))}\right)}{F(\Xi(0))\left(\prod_{k=\Xi(0)+1}^{\infty} \frac{F(k)}{F(k)-F(\Xi(0))}\right)} = 1 - \prod_{k=x+1}^{\infty}\left(1 - \frac{F(\Xi(0))}{F(k)}\right)$$

As in (5.12), we can characterize the tail distribution of T and $\bar{T}(\Xi(0))$. This allows the following asymptotic calculation:

$$\begin{aligned}
&\mathbb{P}(T-\tau(\Xi(0))>s\,|\,T>t)=\frac{\mathbb{P}(\bar{T}(\Xi(0))>s,\,T>t)}{\mathbb{P}(T>t)}=\frac{\mathbb{P}(\bar{T}(\Xi(0))>\max\{t-\tau(\Xi(0)),s\})}{\mathbb{P}(T>t)}\\
&=\frac{1}{\mathbb{P}(T>t)}\int_0^\infty \mathbb{P}(\bar{T}(\Xi(0))>\max\{t-u,s\})F(\Xi(0))e^{-F(\Xi(0))u}du\\
&=\frac{F(\Xi(0))}{\mathbb{P}(T>t)}\left(\int_0^{t-s}\mathbb{P}(\bar{T}(\Xi(0))>t-u)e^{-F(\Xi(0))u}du+\int_{t-s}^\infty \mathbb{P}(\bar{T}(\Xi(0))>s)e^{-F(\Xi(0))u}du\right)\\
&\le\frac{F(\Xi(0))}{\mathbb{P}(T>t)}\left(const.\int_0^{t-s}e^{-F(\Xi(0)+1)(t-u)}e^{-F(\Xi(0))u}du+const.e^{-F(\Xi(0)+1)s}e^{-F(\Xi(0))(t-s)}\right)\\
&\sim\frac{const.}{e^{-F(\Xi(0))t}}\left(e^{-F(\Xi(0)+1)t}\left(1-e^{(F(\Xi(0)+1)-F(\Xi(0)))(t-s)}\right)+e^{-F(\Xi(0)+1)s}e^{-F(\Xi(0))(t-s)}\right)\\
&\sim const.\left(e^{-(F(\Xi(0)+1)-F(\Xi(0)))s}+e^{-(F(\Xi(0)+1)-F(\Xi(0)))s}\right)\quad \text{for } t\to\infty
\end{aligned}$$

Note that *const.* is independent of s and t. □

Example 5.9. For $F(k)=k^\beta$ with $\beta\in\{2,3,\ldots\}$, we have the general formula

$$\prod_{k=x+1}^{\infty}\left(1-\frac{F(\Xi(0))}{F(k)}\right)=\Gamma(x+1)^\beta\prod_{k=0}^{\beta-1}\frac{1}{\Gamma(x+1-\Xi(0)e^{\frac{2\pi ik}{\beta}})}\,,$$

where Γ denotes the Gamma function. Using $\Gamma(k)=(k-1)!$, we get for $\beta=2$:

$$\lim_{t\to\infty}\mathbb{P}(\Xi(t)>x\,|\,T>t)=\frac{\Gamma(x+1)}{\Gamma(x+1+\Xi(0))}=\prod_{k=1}^{\Xi(0)}\frac{1}{x+k}$$

In particular, for $x=\Xi(0)$ we get for the probability that the chain has moved at all

$$\lim_{t\to\infty}\mathbb{P}(\Xi(t)>\Xi(0)\,|\,T>t)=\frac{\Xi(0)!}{(2\Xi(0))!}$$

which vanishes for large initial condition $\Xi(0)\to\infty$. On the other hand, for $\Xi(0)=1$ we have

$$\lim_{t\to\infty}\mathbb{P}(\Xi(t)>x\,|\,T>t)=\frac{1}{x+1}=\frac{1}{2}\quad\text{for } x=\Xi(0)\,.$$

For $\beta\neq 2$, $\beta>1$ and $\Xi(0)=1$ considering the increments yields that the approximation

$$\lim_{t\to\infty}\mathbb{P}(\Xi(t)>x\,|\,T>t)\approx\frac{1}{(\beta-1)(x+1)^{\beta-1}}$$

is accurate even for small x, but not precise. The good precision of this approximation for large t is visible in Figure 5.1.

The discussion in this section does not crucially rely on strict monotonicity of F and we assumed it for simplicity of the presentation and since it is quite natural in many applications. The results could be extended directly to feedback functions with a unique minimum, where the non-exploding chain will spend most of its time.

5.2 Application to generalized Pólya urns

We now apply the findings from Section 5.1 to our generalized Pólya urn $(X(n))_n$ with non-linear feedback functions $F_1, \ldots, F_A$ and $A \geq 2$ agents. We use the definitions and notation introduced in Section 2.1. The connection between Pólya urns and birth processes is provided by the exponential embedding as explained in Section 2.1, where we considered A independent birth processes $(\Xi_i(t))_{t\geq 0}$ with rate functions F_i and sojourn times $\tau_i(k)$, $i \in [A]$. Denoting by $(t_n)_n$ their combined sequence of jump times, we can identify

$$X(n) = (\Xi_1(t_n), \ldots, \Xi_A(t_n)) \quad \text{for all } n \in \mathbb{N}_0 \,.$$

For simplicity of notation, we write

$$T_i := T_i(X_i(0)) = \sup\{t > 0 \colon \Xi_i(t) < \infty\} \in (0, \infty]$$

for the (random) explosion time of Ξ_i and if F_i satisfies (M), denote by

$$g_i(t) := \frac{d}{dt}\mathbb{P}(T_i \leq t)$$

the density of T_i.

5.2.1 Asymptotic size of losing agents

Conditioned on agent i losing, i.e. not to be the monopolist,

$$X_i(\infty) := \lim_{n\to\infty} X_i(n)$$

is a finite random variable, whose tail distribution can be computed using Corollary 5.7. This leads to our next main result, which generalizes previous work in [44, 102, 132] (see Theorem 2.8).

Theorem 5.10. *Assume that at least two agents fulfill (M) and let agent $i \in [A]$ be one of them. Then*

$$\mathbb{P}\left(X_i(\infty) = x\right) \sim \frac{\mathbb{E}g_i\left(\min_{j\neq i} T_j\right)}{F_i(x)} \quad \text{as } x \to \infty$$

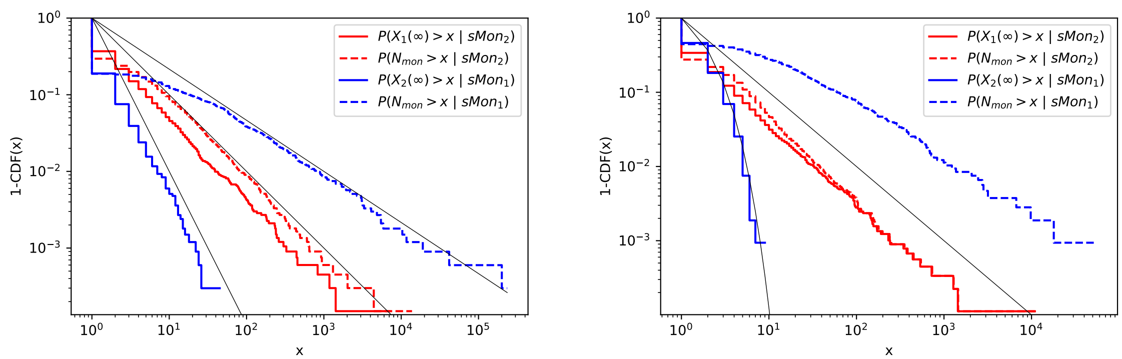
and consequently

$$P(X_i(\infty) > x \,|\, sMon_i^c) \sim \frac{\mathbb{E}g_i\left(\min_{j\neq i} T_j\right)}{\mathbb{P}(T_i > \min_{j\neq i} T_j)} \sum_{k=x+1}^{\infty} \frac{1}{F_i(k)} \quad \text{as } x \to \infty \,.$$

Proof. This is an immediate consequence of Corollary 5.7 since $S := \min_{j\neq i} T_j$ is independent of $(\Xi_i(t))_{t\geq 0}$ and

$$\{X_i(\infty) = x\} = \{\Xi_i(S) = x\} \quad \text{as well as} \quad sMon_i^c = \{T_i > S\} \,.$$

Note that $\mathbb{E}g_i\left(\min_{j\neq i} T_j\right) \in (0, \infty)$ as g_i is bounded (Lemma 5.2) and at least one T_j is finite. □



(a) $F_1(k) = k^2$, $F_2(k) = k^3$ (b) $F_1(k) = k^2$, $F_2(k) = e^k$

Figure 5.3: Empirical distribution of $X_i(\infty)$ and N_{mon} (defined in (5.18)) for $A = 2$ and $X_1(0) = X_2(0) = 1$. The black lines show the tail decay predicted by Theorem 5.10 and Corollary E.1. Note that these predicted tails do only claim exactness up to constants, i.e. up to a parallel shift in the plot. 10,000 simulations were executed each.

Figure 5.3 illustrates Theorem 5.10 for polynomial feedback. Remarkably, the tail of $X_i(\infty)$ conditioned on $sMon_i^c$ does not depend on the feedback functions of other agents up to a constant prefactor. In particular, $X_i(\infty)$ has a power-law tail even if there is another agent with exponential feedback. For the existence of moments of $X_i(\infty)$ conditioned on $sMon_i^c$, criterion (5.9) holds analogously. As explained in Example 5.5, $\mathbb{P}(X_i(\infty) > x \,|\, sMon_i^c)$ has a power-law tail for polynomially increasing F_i, whereas it has exponential tail for exponentially increasing F_i. For $F_i(k) = k(\log k)^\beta$ with $\beta \geq 1$, the tail decay is even logarithmically slow.

An astonishing consequence of Theorem 5.10 is the **"loser paradox"** explained in the following.

1. Theorem 5.10 implies that the tail of $X_i(\infty)$ conditioned on $sMon_i^c$ is heavier the weaker the feedback is. This means that among agents with feedback satisfying (M), losers with slowly increasing F_i are more likely to win in many steps than other losers with stronger feedback. In analogy to the same phenomenon for the single birth process, if an agent with strong feedback wins in many steps, then they are likely to be the monopolist. On the other hand, if this agent is not the monopolist, then they are likely to have won only a few steps.

2. Let us now consider a system, where at least one agent fulfills (M), but not agent i. Then the process exhibits strong monopoly, but $\mathbb{P}(sMon_i) = 0$ and hence $\mathbb{P}(X_i(\infty) < \infty) = 1$ and agent i loses. Since such agents are not affected by conditioning on $sMon_i^c$, via canonical coupling it is obvious that $X_i(\infty)$ is stochastically larger the faster F_i increases. So among agents with feedback that does not satisfy (M) those with stronger feedback are likely to win more steps.

3. Summarizing both cases, the tail of $X_i(\infty)$ for a losing agent (possibly conditioning on $sMon_i^c$) is heavier for feedback functions closer to the transition given by (M). For a better understanding of this tail-maximising phenomenon, a detailed discussion of the tail behaviour close to the transition (M) is provided at the end of Section 5.3.

In general, the prefactor $\mathbb{E}g_i\left(\min_{j\neq i} T_j\right)$ in Theorem 5.10 is decreasing in A, i.e. the probability that a loser wins many steps is smaller in larger systems. Moreover, this prefactor vanishes for $A \to \infty$, if infinitely many agents satisfy (M). This complies with the idea that in large systems a typical agent does not even win a single step (see Appendix D.1). Surprisingly, large systems behave totally different if only one agent satisfies (M), such that this agent is almost surely the monopolist. In this case, the distribution of $X_i(\infty)$ is independent of the number of other agents as a direct consequence of the exponential embedding. Hence, in a world with super-linear (polynomial) feedback and many agents, there will be a few agents with high wealth, but the vast majority will not win anything, whereas in a world with one dominant agent (the super-linear one) all sub-linear losers have good chances to win some steps.

5.2.2 Correlations among losing agents

Next, we have a closer look at the dependencies of several losing agents. Note that the correlation of $X_i(\infty)$ and $X_j(\infty)$ on the conditioned space $sMon_i^c \cap sMon_j^c$ does in general not exist due to infinite second moments, e.g. for $F_i(k) = k^\beta$ with $\beta \in (1, 3]$.

Theorem 5.11. *Let $a \in [A-1]$. Assume that at least $a+1$ agents satisfy (M) and that agents $1, \dots, a$ are among them. Set $S \coloneqq \min\{T_i : i = a+1, \dots, A\}$ and $S_i \coloneqq \min\{T_j : j \in [A]\backslash\{i\}\}$. Then:*

$$\mathbb{P}\left(X_1(\infty) = x_1, \dots, X_a(\infty) = x_a\right) \sim \frac{\mathbb{E}\left[\prod_{i=1}^a g_i(S)\right]}{\prod_{i=1}^a \mathbb{E}g_i(S_i)} \cdot \prod_{i=1}^a \mathbb{P}\left(X_i(\infty) = x_i\right) \quad \text{for } x_1, \dots, x_a \to \infty$$

Proof. Denote by $\mathbb{P}_S$ the law of S. Using the independence property of the exponential embedding, we get:

$$\begin{aligned}
&\mathbb{P}\left(X_1(\infty) = x_1, \dots, X_a(\infty) = x_a\right) = \mathbb{P}\left(\Xi_1(S) = x_1, \dots, \Xi_a(S) = x_a\right) \\
&= \int_0^\infty \mathbb{P}\left(\Xi_1(s) = x_1, \dots, \Xi_a(s) = x_a\right) \, d\mathbb{P}_S(s) = \int_0^\infty \prod_{i=1}^a \mathbb{P}\left(\Xi_i(s) = x_i\right) \, d\mathbb{P}_S(s) \\
&\sim \int_0^\infty \prod_{i=1}^a \frac{g_i(s)}{F_i(x_i)} \, d\mathbb{P}_S(s) = \mathbb{E}\left[\prod_{i=1}^a g_i(S)\right] \prod_{i=1}^a \frac{1}{F_i(x_i)} \quad \text{for } x_1, \dots, x_a \to \infty
\end{aligned}$$

In the last line, we apply Theorem 5.1 and Corollary 5.6 in the same manner as in the proof of Corollary 5.7. The claim follows then via Theorem 5.10. □

In particular, for polynomial feedback, $(X_1(\infty), X_2(\infty))$ has a two-dimensional heavy-tailed distribution on $sMon_1^c \cap sMon_2^c$ in the sense of [87]. Denote by

$$\mathbb{P}_{A,a}(\cdot) \coloneqq \frac{\mathbb{P}((\cdot) \cap sMon_1^c \cap \dots sMon_a^c)}{\mathbb{P}(sMon_1^c \cap \dots sMon_a^c)}$$

the conditioned probability measure on $sMon_1^c \cap \ldots \cap sMon_a^c$. Then we can rephrase Theorem 5.11 as

$$\mathbb{P}_{A,a}\left(X_1(\infty) > x_1, \ldots, X_a(\infty) > x_a\right) \sim c(A,a) \cdot \prod_{i=1}^{a} \mathbb{P}_{A,a}\left(X_i(\infty) > x_i\right) \quad \text{for } x_1, \ldots, x_a \to \infty\,, \tag{5.14}$$

where

$$c(A,a) \coloneqq \frac{\mathbb{E}\left[\prod_{i=1}^{a} g_i(S)\right]}{\prod_{i=1}^{a} \mathbb{E}\left[\left(\prod_{j\in[a]\setminus\{i\}} G_j(S)\right) g_i(S)\right]} \cdot \mathbb{P}\left(sMon_1^c \cap \ldots \cap sMon_a^c\right)^{a-1}\,. \tag{5.15}$$

and $G_i(s) \coloneqq \mathbb{P}(T_i > s)$, since for $x_i \to \infty$

$$\begin{aligned}
&\mathbb{P}_{A,a}\left(X_i(\infty) = x_i\right) = \frac{\mathbb{P}(\Xi_i(S) = x_i,\, \forall j \in [a]\colon T_j > S)}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} && (5.16)\\
&= \frac{1}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} \int_0^\infty \mathbb{P}(\Xi_i(s) = x_i,\, \forall j \in [a]\colon T_j > s)\, d\mathbb{P}_S(s)\\
&= \frac{1}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} \int_0^\infty \mathbb{P}(\Xi_i(s) = x_i) \prod_{j\in[a]\setminus\{i\}} G_j(s)\, d\mathbb{P}_S(s)\\
&\sim \frac{1}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} \int_0^\infty \frac{g_i(s)}{F_i(x_i)} \prod_{j\in[a]\setminus\{i\}} G_j(s)\, d\mathbb{P}_S(s) \quad \text{for } x \to \infty\\
&= \frac{\mathbb{E}\left[\left(\prod_{j\in[a]\setminus\{i\}} G_j(S)\right) g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} \cdot \frac{1}{F_i(x_i)} \sim \frac{\mathbb{E}\left[\left(\prod_{j\in[a]\setminus\{i\}} G_j(S)\right) g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)\mathbb{E}g_i(S_i)} \cdot \mathbb{P}(X_i(\infty) = x_i)\\
&= \frac{\mathbb{E}\left[\left(\prod_{j\in[a]\setminus\{i\}} G_j(S)\right) g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)\mathbb{E}g_i(S_i)} \cdot \mathbb{P}(sMon_i^c) \cdot \mathbb{P}(X_i(\infty) = x_i \mid sMon_i^c)
\end{aligned}$$

i.e. the tail of $X_i(\infty)$ conditioned on $sMon_i^c$ is equal to its tail conditioned on $sMon_1^c \cap \ldots \cap sMon_a^c$ up to a constant. This constant is necessarily one in a fully symmetric situation, since the distribution of $X_i(\infty)$ on $sMon_i^c$ is then independent of who the monopolist is. In a non-symmetric situation, the information who won contains information on the explosion time S, which affects the distribution of $X_i(\infty)$. For example, if an agent with strong feedback is the monopolist, then S is likely to be small, such that $X_i(\infty)$ is rather small, too. Hence, when $F_1 = \ldots = F_A$ and $X_1(0) = \ldots = X_A(0)$, then we can simply write

$$c(A,a) = \frac{\mathbb{E}\left[g(S)^a\right]}{\left(\mathbb{E}g(S)\right)^a} \cdot \frac{\prod_{i=1}^{a} \mathbb{P}(sMon_i^c)}{\mathbb{P}(sMon_1^c \cap \ldots sMon_a^c)} = \frac{\mathbb{E}\left[g(S)^a\right]}{\left(\mathbb{E}g(S)\right)^a} \cdot \frac{(A-1)^a}{A^{a-1}(A-a)}\,,$$

where $g = g_i$ and $G = G_i$. The first factor is larger than one due to Jensen's inequality and the second factor is larger than one due to

$$\frac{(A-1)^a}{A^{a-1}(A-a)} > 1 \quad \Leftrightarrow \quad \left(1 - \frac{1}{A}\right)^a > 1 - \frac{a}{A} \tag{5.17}$$

and Bernoulli's inequality. Hence, we have $c(A,a) > 1$ in the symmetric case.

In general, the dependence between several losers can be considered as weak since the tail weight of a loser does not change when the result of other losers is known. The remaining tail dependence is captured by the constant $c(A,a)$, which we shortly discuss now. Heuristically, the event that a loser wins in many steps can occur in two ways. First, it is possible that the underlying birth process of this agent increases fast. This effect is independent between several losers. Second, it is possible that the explosion time of the winner is relatively late, such that all losers are more likely to win in many steps. This second effect accounts for a slightly **positive correlation of losers**, corresponding to $c(A,a) > 1$. The following proposition formally substantiates this idea also for non-diverging x_i.

Proposition 5.12. *Assume $A \geq 3$, $X_1(0) = \ldots = X_A(0)$ and $F_1 = \ldots = F_A = F$, where F satisfies (M). Then we have $c(A,a) > 1$ and*

$$\mathbb{P}_{A,a}\left(X_1(\infty) > x, \ldots, X_a(\infty) > x\right) > \prod_{i=1}^{a} \mathbb{P}_{A,a}\left(X_i(\infty) > x\right)$$

for any $x \geq X_1(0)$ and $a \in [A-1]$.

Proof. $c(A,a) > 1$ has been derived above. Due to the assumed symmetry, we can define $h(s) := \mathbb{P}(\Xi_i(s) > x)$. Denote by $\mathbb{P}_S$ the law of S (defined in Theorem 5.11). Using Jensen's inequality, we get:

$$\begin{aligned}
\mathbb{P}_{A,a}\left(X_1(\infty) > x, \ldots, X_a(\infty) > x\right) &= \frac{\mathbb{P}\left(\Xi_1(S) > x, \ldots, \Xi_a(S) > x\right)}{\mathbb{P}(sMon_1^c \cap \ldots \cap sMon_a^c)} \\
&= \frac{1}{\mathbb{P}(sMon_1^c \cap \ldots \cap sMon_a^c)} \int_0^\infty \mathbb{P}\left(\Xi_1(s) > x, \ldots, \Xi_a(s) > x\right) d\mathbb{P}_S(s) \\
&= \frac{A}{A-a} \int_0^\infty h(s)^a \, d\mathbb{P}_S(s) > \frac{A}{A-a} \left(\int_0^\infty h(s) \, d\mathbb{P}_S(s) \right)^a \\
&= \frac{A}{A-a} \mathbb{P}(\Xi_1(S) > x)^a > \frac{A}{A-a} \mathbb{P}(X_1(\infty) > x, \, sMon_1^c)^a \\
&= \frac{A}{A-a} \left(\mathbb{P}(X_1(\infty) > x, \, sMon_1^c, \ldots, sMon_a^c) \cdot \frac{A-1}{A-a} \right)^a \\
&= \frac{A(A-1)^a}{(A-a)^{a+1}} \prod_{i=1}^{a} \left[\mathbb{P}_{A,a}\left(X_i(\infty) > x\right) \mathbb{P}(sMon_1^c, \ldots, sMon_a^c) \right] \\
&= \frac{A(A-1)^a}{(A-a)^{a+1}} \cdot \left(\frac{A-a}{A} \right)^a \prod_{i=1}^{a} \mathbb{P}_{A,a}\left(X_i(\infty) > x\right) > \prod_{i=1}^{a} \mathbb{P}_{A,a}\left(X_i(\infty) > x\right)
\end{aligned}$$

The last step follows again via Bernoulli's inequality. Throughout the calculation, we exploited the supposed symmetry of all agents. ☐

Appendix D.3 examines the dependence of $c(A,a)$ on the system size A. In Appendix D.1 (in particular Theorem D.1) we will first see that for large A, a randomly chosen agent does not win a single step with high probability. Recall that $c(A,a)$ only measures the tail dependence of agents, i.e. we basically only consider agents who won many steps. Surprisingly, $c(A,a)$ turns out to be increasing with A, i.e. the dependence between losers in the sense captured by $c(A,a)$ is stronger in large systems. This seems to be opposed to the linear case, where we observe asymptotic independence for $A \to \infty$ (Theorem D.6), but we also have a trivial asymptotic independence in the super-linear case (Corollary D.3) since typical losers do not win anything. Moreover, $c(A,a)$ is increasing in a since the information that several losers won many steps is a stronger hint for a late explosion time of the winner, such that the considered agent is more likely to win many steps, too.

5.2.3 Further consequences of Theorem 5.10 and Theorem 5.11

Theorem 5.10 and Theorem 5.11 are useful to determine the tail behaviour of further random variables arsing in the context of our generalized Pólya urn model. This subsection presents some examples.

Corollary 5.13. *In the situation of Theorem 5.11, the following asymptotics hold for $x \to \infty$.*

1.

$$\mathbb{P}_{A,a}(\min(X_1(\infty),\dots,X_a(\infty)) > x) \sim \frac{\mathbb{E}\left[\prod_{i=1}^a g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \dots \cap sMon_a^c)} \cdot \prod_{i=1}^a \left(\sum_{k=x+1}^{\infty} \frac{1}{F_i(k)}\right)$$

2.

$$\mathbb{P}_{A,a}(\max(X_1(\infty),\dots,X_a(\infty)) > x) \sim \sum_{i=1}^a \left(\frac{\mathbb{E}\left[\left(\prod_{j\neq i} G_j(S)\right) g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \dots \cap sMon_a^c)} \sum_{k=x+1}^{\infty} \frac{1}{F_i(k)}\right)$$

3. *If $F_1,\dots,F_a$ are regular varying, then:*

$$\mathbb{P}_{A,a}(X_1(\infty) + \dots + X_a(\infty) > x) \sim \sum_{i=1}^a \left(\frac{\mathbb{E}\left[\left(\prod_{j\neq i} G_j(S)\right) g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \dots \cap sMon_a^c)} \sum_{k=x+1}^{\infty} \frac{1}{F_i(k)}\right)$$

Proof. 1. We rephrase Theorem 5.11 again as

$$\mathbb{P}_{A,a}(X_1(\infty) > x_1,\dots,X_a(\infty) > x_a) \sim \frac{\mathbb{E}\left[\prod_{i=1}^a g_i(S)\right]}{\mathbb{P}(sMon_1^c \cap \dots \cap sMon_a^c)} \cdot \prod_{i=1}^a \left(\sum_{k=x_i+1}^{\infty} \frac{1}{F_i(k)}\right).$$

Then the claim follows from

$$\mathbb{P}_{A,a}(\min(X_1(\infty),\dots,X_a(\infty)) > x) = \mathbb{P}_{A,a}\left(X_1(\infty) > x, \dots, X_a(\infty) > x\right).$$

2. This follows from (5.16) and the inclusion-exclusion formula via

$$\mathbb{P}_{A,a}(\max(X_1(\infty),\ldots,X_a(\infty))>x)=\mathbb{P}_{A,a}\left(\bigcup_{i=1}^{a}\{X_i(\infty)>x\}\right)\sim\sum_{i=1}^{a}\mathbb{P}_{A,a}(X_i(\infty)>x)\,.$$

3. Karamata's Theorem implies that $x\mapsto\sum_{k=x+1}^{\infty}\frac{1}{F_i(k)}$ are also regularly varying for $i=1,\ldots,\ a$. Then the claim follows from (5.16) and Lemma 5.14 applied on the conditioned probability space $sMon_1^c\cap\ldots\cap sMon_a^c$. ☐

Lemma 5.14. *Let $X_1,\ldots,X_k$ be random variables, such that $x\mapsto\mathbb{P}(X_i>x)$ is regular varying for all $i=1,\ldots,k$. Moreover, assume that the X_i are at most weakly dependent in the sense that*

$$\forall i,j\in[k]\colon\ \lim_{x\to\infty}\frac{\mathbb{P}(X_i>x)}{\mathbb{P}\left(X_i>x,\,X_j>x\right)}=\infty\,.$$

Then:

$$\mathbb{P}(X_1+\ldots+X_k>x)\sim\mathbb{P}(X_1>x)+\ldots+\mathbb{P}(X_k>x)\quad\textit{for }x\to\infty$$

Proof. We only prove the case $k=2$ since general k follows directly by induction. Then we have the following lower bound:

$$\mathbb{P}(X_1+X_2>x)\geq\mathbb{P}(X_1>x)+\mathbb{P}(X_2>x)-\mathbb{P}(X_1>x,\,X_2>x)\sim\mathbb{P}(X_1>x)+\mathbb{P}(X_2>x)$$

For an upper bound, take $0<\delta<\frac{1}{2}$. Then

$$\begin{aligned}\mathbb{P}(X_1+X_2>x)&\leq\mathbb{P}(X_1>(1-\delta)x)+\mathbb{P}(X_2>(1-\delta)x)+\mathbb{P}(X_1>\delta x,\,X_2>\delta x)\\&\sim(1-\delta)^{\alpha_1}\mathbb{P}(X_1>x)+(1-\delta)^{\alpha_2}\mathbb{P}(X_2>x)\,,\end{aligned}$$

where α_i is the index of regular variation of $x\mapsto\mathbb{P}(X_i>x)$. $\delta\to0$ implies the claim. ☐

Note that in Corollary 5.13 the maximum and sum have the same tail. This is a characterizing property of sub-exponential random variables and, hence, the assumption of regular varying feedback cannot be omitted in 3. of Corollary 5.13.

Again, the tail decay of the sum, minimum and maximum of a losers only depends on $F_1,\ldots,F_a$, but not on $F_{a+1},\ldots,F_A$ (up to the constant prefactor). Corollary 5.13 also contains detailed information about the sum, minimum and maximum of all losers. To see that, denote by $W\in[A]$ the random winner of the process and write

$$\mathbb{P}\left(\sum_{j\neq W}X_j(\infty)>x\right)=\sum_{i=1}^{A}\mathbb{P}(sMon_i)\mathbb{P}\left(\sum_{j\neq i}X_j(\infty)>x\,\middle|\,\bigcap_{j\neq i}sMon_j^c\right)$$

and analogously for the maximum and minimum. Let us finally discuss this in the polynomial case.

Example 5.15. Let $F_i(k) = \alpha_i k^{\beta_i}$ for $\alpha_i > 0$ and $1 < \beta_1 \leq \ldots \leq \beta_A$. Then

$$\mathbb{P}\left(\sum_{i \neq W} X_i(\infty) > x\right) \sim \mathbb{P}\left(\max_{i \neq W} X_i(\infty) > x\right) \sim const.x^{1-\beta_1} \quad \text{for } x \to \infty$$

and

$$\mathbb{P}\left(\min_{i \neq W} X_i(\infty) > x\right) \sim const.x^{A-1-\beta_1-\ldots-\beta_{A-1}} \quad \text{for } x \to \infty\,.$$

In both cases, the tail decay is independent of β_A up to a constant. Moreover, if $\beta_1 < \beta_2$ and the losers win many steps, then it is most likely that in fact the loser with the heaviest tail and the weakest feedback wins most of these steps, i.e.

$$\mathbb{P}\left(X_1(\infty) > x \,\middle|\, \sum_{i=1}^{a} X_i(\infty) > x, sMon_1^c \cap \ldots \cap sMon_a^c\right) \xrightarrow{x \to \infty} 1$$

for any $a < A$, consistent with the idea of Sanov's Theorem.

Moreover, Theorem 5.10 can be used to generalize previous results from [132, 44] concerning the **time of monopoly**, which is defined as

$$N_{mon} \coloneqq \min\{n \geq 1 \colon \exists i \in [A]\, \forall\, m \geq n \colon X(m) - X(m-1) = e^{(i)}\}\,, \tag{5.18}$$

i.e. the index of the last step won by a loser (plus one). Using Theorem 5.10, we can calculate the tail distribution of N_{mon} for polynomial feedback $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_i > 1$ and $\alpha_i > 0$, $i \in [A]$. Assume $\beta_2 \leq \beta_j$ for all $j \geq 2$, i.e. among the losers agent 2 has the weakest feedback. Then the following can be shown (under mild further assumptions), conditional on the first agent being the winner:

1. If $\beta_2 \geq \beta_1 - 1$, then we have

$$\mathbb{P}(N_{mon} > n \,|\, sMon_1) \asymp n^{-(\beta_2-1)(\beta_1-1)/\beta_2} \quad \text{for } n \to \infty\,. \tag{5.19}$$

2. If $1 < \beta_2 \leq \beta_1 - 1$, then we have

$$\mathbb{P}(N_{mon} > n \,|\, sMon_1) \asymp n^{-(\beta_2-1)} \quad \text{for } n \to \infty\,. \tag{5.20}$$

A proof is presented in Appendix E.1 as well as a discussion of exponential feedback. As opposed to the wealth of losers, the tail of the time of monopoly depends also on the feedback of the winner. Note that the tail behaviour of N_{mon} is independent of the losers with stronger feedback F_j, $j \geq 3$. Figure 5.4 illustrates this situation for $A = 2$. This result grants some new interesting insights in the large deviation behaviour of asymmetric generalized Pólya urns. For example, when the monopoly sets in late, this happens in different ways depending on the feedback. For simplicity, we assume $A = 2$ and condition on agent 1 to be the winner. The following is established rigorously in Corollary E.4:

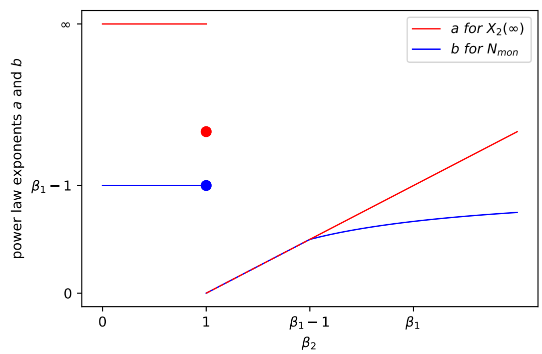


Figure 5.4: The power law exponents $a, b > 0$ of $X_1(\infty)$ and N_{mon} for fixed feedback of the winner $F_1(k) = k^{\beta_1}$, $\beta_1 > 2$, and varying feedback of the loser $F_2(k) = k^{\beta_2}$, $\beta_2 > 0$, i.e. $A = 2$, $\mathbb{P}(X_2(\infty) > x \,|\, sMon_1) \asymp x^{-a}$ for $x \to \infty$ and $\mathbb{P}(N_{mon} > n \,|\, sMon_1) \asymp x^{-b}$ for $n \to \infty$. $a = \infty$ corresponds to tails lighter than a power law. Theorem 5.10 and Corollary E.1 cover the case $\beta_2 > 1$, and Theorem 5.16 and Corollary E.6 cover the case $\beta_2 \leq 1$. The transition at $\beta_2 = 1$ is also discussed in Section 5.3. Note that $a \geq b$ since on the event $sMon_1$ we have $N_{mon} \geq X_2(\infty)$.

1. If the feedback of the winner is significantly stronger than the feedback of the loser (i.e. $1 < \beta_2 \leq \beta_1 - 1$), then the tail of N_{mon} is independent of β_1 and equals the one of the wealth of the loser $X_2(\infty)$ up to a constant prefactor ($a = b$ in Figure 5.4). So the loser wins most steps at the beginning (even more than the winner), when suddenly the strong monopoly sets in and the loser does not win any further steps.

2. If the feedback of the winner is at most slightly stronger than the feedback of the loser (i.e. $\beta_2 \geq \beta_1 - 1$), then the tail of N_{mon} is heavier than the one of $X_2(\infty)$. Hence, the loser can still win some late steps when the advantage of the winner is already large ($a > b$ in Figure 5.4).

3. At the point of transition $\beta_2 = \beta_1 - 1$, we observe an intermediate regime, where the process is balanced at the beginning, when suddenly the strong monopoly sets in and the monopolist wins all further steps.

Note that this reflects our findings on total monopoly for polynomial feedback (Example 3.9).

5.3 Sub-linear agents

So far, we mostly considered agents satisfying the monopoly condition (M). Nevertheless, any system with at least one agent fulfilling (M) exhibits strong monopoly, even if some agents do not satisfy (M). The latter agents are almost surely not the monopolist and hence, $X_i(\infty)$ is a finite random variable for them. The following theorem characterizes the corresponding tail distribution.

Theorem 5.16. *Assume that the set* $M := \{i \in [A] : F_i$ *satisfies* $(M)\}$ *of possible monopolists is non-empty but* $1 \notin M$. *Moreover, assume* $\lim_{k\to\infty} F_1(k) = \infty$ *and that all* F_i, $i \in M$ *are strictly monotone. Then*

$$-\log(\mathbb{P}(X_1(\infty) > x)) \sim d \sum_{k=X_1(0)}^{x} \frac{1}{F_1(k)} \quad \text{for } x \to \infty\,, \quad \text{where} \quad d := \sum_{i\in M} F_i(X_i(0))\ .$$

Proof. Let $T := \min_{i\in M} T_i$ be the smallest explosion time of the possible monopolists in the exponential embedding. Moreover, write $S(x) := \sum_{k=X(0)}^{x} \tau_1(k)$ and denote by $\mu_{S(x)} = \mathbb{P}[S(x) \in \cdot]$ the law of $S(x)$. Using the independence of $S(x)$ and T, we get:

$$\begin{aligned}\mathbb{P}(X_1(\infty) > x) = \mathbb{P}(\Xi_1(T) > x) = \mathbb{P}\left(S(x) < T\right) &= \int_0^\infty \mathbb{P}(s < T)\, d\mu_{S(x)}(s) \\ &= \int_0^\infty \prod_{i\in M} \mathbb{P}(T_i > s)\, d\mu_{S(x)}(s)\end{aligned}$$

Next, (5.12) implies

$$e^{-F_i(X_i(0))s} = \mathbb{P}(\tau_i(X(0)) > s) < \mathbb{P}(T_i > s) < ce^{-F_i(X_i(0))s} \tag{5.21}$$

for a constant $c > 0$. Hence:

$$\mathbb{E}\left[e^{-dS(x)}\right] < \mathbb{P}(X_1(\infty) > x) < c^{\#M}\mathbb{E}\left[e^{-dS(x)}\right]$$

Then the following Lemma 5.17 and

$$\lim_{k\to\infty} F_1(k) = \infty \quad \Longrightarrow \quad \lim_{k\to\infty} \frac{\sum_{l=1}^{k} \frac{1}{F_1(k)}}{\sum_{l=1}^{k} \frac{1}{F_1(k)^2}} = \infty$$

finally imply the claim. □

Note that estimate (5.21) does in general not hold if there was some $k > X_i(0)$ such that $F_i(k) = F_i(X_i(0))$. Moreover, Theorem 5.16 does not hold for non-diverging feedback, since e.g. for constant $F_i(k) = 1$ we have

$$\mathbb{E}\left[e^{-dS(x)}\right] = \exp\left(-\log(1+d) \sum_{k=X_i(0)}^{x} \frac{1}{F_i(x)}\right) = (1+d)^{X_i(0)-x-1}\,.$$

Lemma 5.17. *Assume that* $X_1, \ldots, X_k$ *are independent random variables, which are exponentially distributed with parameters* $\lambda_1, \ldots, \lambda_k > 0$. *The moment generating function of* $S := X_1 + \ldots + X_k$ *is bounded by*

$$\exp\left\{-s\sum_{l=1}^{k}\frac{1}{\lambda_k}\right\} < \mathbb{E}\left[e^{-sS}\right] < \exp\left\{-s\sum_{l=1}^{k}\frac{1}{\lambda_k} + s^2\sum_{l=1}^{k}\frac{1}{\lambda_k^2}\right\}$$

for all $s > 0$.

Proof. Given $\mathbb{E}[e^{-sX_l}] = \frac{\lambda_l}{\lambda_l+s}$, we calculate:

$$\mathbb{E}\left[e^{-sS}\right] = \prod_{l=1}^{k} \mathbb{E}^{-sX_l} = \prod_{l=1}^{k} \frac{\lambda_l}{\lambda_l + s} = \exp\left\{\sum_{l=1}^{k} (\log(\lambda_l) - \log(\lambda_l + s))\right\}$$

Moreover, since $\frac{d}{ds}\log(s) = \frac{1}{s}$, we can estimate

$$\log(\lambda_l + s) - \log(\lambda_l) \leq \frac{s}{\lambda_s},$$

and analogously

$$\log(\lambda_l + s) - \log(\lambda_l) \geq \frac{s}{\lambda_l + s} = \frac{s}{\lambda_l} + s\left(\frac{1}{\lambda_l + s} - \frac{1}{\lambda_l}\right) \geq \frac{s}{\lambda_l} - \frac{s^2}{\lambda_l^2},$$

which completes the proof. □

A closer look at the proof reveals the following refinement of Theorem 5.16:

$$0 < \log(\mathbb{P}(X_1(\infty) > x)) + d \sum_{k=X_1(0)}^{x} \frac{1}{F_1(k)} < d^2 \sum_{k=X_1(0)}^{x} \frac{1}{F_1(k)^2} + const. \quad \text{for all } x \geq X_1(0)$$

Let us discuss some interesting examples.

Example 5.18. In the situation of Theorem 5.16, we consider various choices of the sub-linear feedback function F_1.

1. Let $F_1(k) = k \log k$. Then

$$\mathbb{P}(X_1(\infty) > x) \asymp (\log x)^{-d} \quad \text{for } x \to \infty. \tag{5.22}$$

Hence, the tail decay is logarithmic and depends on d.

2. Let $F_1(k) = k(\log k)^\beta$ for $\beta < 1$. Then

$$\mathbb{P}(X_1(\infty) > x) \asymp e^{-\frac{d}{1-\beta}(\log x)^{1-\beta}} \quad \text{for } x \to \infty. \tag{5.23}$$

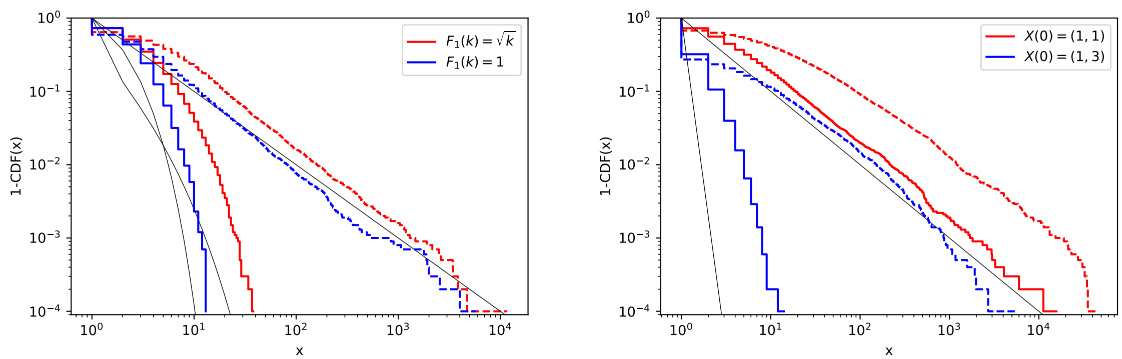
Hence, $X_1(\infty)$ has a heavy-tailed distribution, which is heavier for larger β. In particular, for the linear case $\beta = 0$ we find a power law distribution with exponent depending on β and the initial values included in d, i.e.

$$\mathbb{P}(X_1(\infty) > x) \asymp x^{-d} \quad \text{for } x \to \infty. \tag{5.24}$$

3. Let $F_1(k) = k^\beta$ for $\frac{1}{2} < \beta < 1$. Then

$$\mathbb{P}(X_1(\infty) > x) \asymp e^{-\frac{d}{1-\beta}x^{1-\beta}} \quad \text{for } x \to \infty.$$

Hence, the tail of $X_1(\infty)$ is lighter than a power law, but heavier than an exponential. Moreover, the tail is lighter the smaller β is.



(a) $X(0) = (1, 1)$ and varying F_1

(b) $F_1(k) = k$ and varying $X(0)$

Figure 5.5: Empirical distribution of $X_1(\infty)$ (full line) and N_{mon} (dotted line) for $A = 2$, $F_2(k) = k^2$ and different F_1 and $X(0)$. The black lines show the tail decay predicted by Theorem 5.16 and Corollary E.6. Note that these predicted tails are only exact up to constants, i.e. up to a parallel shift in the plot. 10,000 simulations were executed each.

4. Let $F_1(k) = k^{\frac{1}{2}}$. Then

$$e^{-2d\,x^{1/2}} \prec \mathbb{P}(X_1(\infty) > x) \prec e^{-2d\,x^{1/2}} x^{d^2} \quad \text{for } x \to \infty\,.$$

   The same observation as in 3. holds here.

5. Let $F_1(k) = k^\beta$ for $0 < \beta < \frac{1}{2}$. Then

$$e^{-\frac{d}{1-\beta}x^{1-\beta}} \prec \mathbb{P}(X_1(\infty) > x) \prec e^{-\frac{d}{1-\beta}x^{1-\beta} + \frac{d^2}{1-2\beta}x^{1-2\beta}} \quad \text{for } x \to \infty$$

   consistent with 3. and 4..

6. Let $F_1(k) = \lambda > 0$ be constant. Then Theorem 5.16 is not applicable, but Ξ_1 is a homogeneous Poisson process. Hence, $X_1(\infty)$ has a Poisson distribution.

Remarkably, the tail decay of $X_i(\infty)$ depends significantly on initial values of super-linear agents for feedback F_1 close to the transition (M), in contrast to super-linear F_1 in Theorem 5.10. Note that in Theorem 5.16 there is no dependence on the precise feedback function of other agents. Moreover, the distribution of $X_i(\infty)$ does not depend on the feedback of other sub-linear agents due to the exponential embedding. These findings are illustrated by the simulation shown in Figure 5.5. As mentioned in (5.19) and (5.20) and explained in Appendix E.2, the time when the monopoly occurs in the situation of this section has a power-law distribution (for polynomial feedback) and the exponent depends on the feedback of the winner.

By combining Theorem 5.10 and Theorem 5.16, we gain a precise understanding of the "loser paradox" mentioned in Section 5.2. This seeming paradox is that the tail of $X_i(\infty)$

(if necessary conditioned on $sMon_i^c$) is heaviest, when the feedback of agent i is close to the transition (M). This observation is quantified by Example 5.18 and Example 5.5. At the point of **transition for almost linear feedback**, these two examples are consistent. To see that, consider again $F_i(k) = k(\log k)^\beta$ and condition on agent i to be a loser. Recall that throughout this chapter we assumed that at least one other agent satisfies the monopoly condition (M).

1. For $\beta \in (-\infty, 0)$, the tail of $X_i(\infty)$ is heavier than exponential, but lighter than a power law, see (5.23). This does also hold for $F_i(k) = k^{\beta'}$ with $\beta' \in (0, 1)$ (3. to 5. in Example 5.18).

2. In the linear case $\beta = 0$, we have a power law tail x^{-d} with exponent depending on initial values of the super-linear agents, see (5.24).

3. For $\beta \in (0, 1)$, the tail of $X_i(\infty)$ is even heavier than a power law, but the decay is faster than logarithmic, see (5.23). The tail weight is increasing in $\beta < 1$ and depends on the initial values.

4. For $\beta \geq 1$, the tail of $X_i(\infty)$ decays logarithmically, i.e. $\mathbb{P}(X_1(\infty) > x) \asymp (\log x)^{-d}$ for $\beta = 1$, where d depends on initial values, and $\mathbb{P}(X_1(\infty) > x) \asymp (\log x)^{1-\beta}$ for $\beta > 1$, see (5.22) and (5.11). The tail weight is now decreasing in $\beta > 1$, which is consistent with the power law distribution of $X_i(\infty)$ for $F_i(k) = k^{\beta'}$, $\beta' > 1$, see (5.10).

Chapter 6

Wages and Capital Returns in a Pólya Urn Model

It is a widely observed phenomenon that wealth is distributed significantly more unequal than wages. In this chapter, we study this phenomenon using a new extension of Pólya's urn, modelling wealth growth through wages and capital returns. In Section 6.2, we formally introduce the model and provide a mathematical analysis. In Section 6.3, we discuss the cases of $A = 2$ and $A = 3$ agents in order to gain a visual understanding of the different regimes of the process. In Section 6.4, we fit the parameters from real-world data in Germany, so that simulation results reproduce the empirical wealth distribution and recent dynamics in Germany quite accurately. This allows interesting predictions for future developments and insights on the importance of wages and capital returns for wealth aggregation. In Section 6.5, we examine within our model if unequal investment skills pose an alternative explanation for the gap between wealth and wage distribution. Finally, in Section 6.6, we bring together our numerical and theoretical findings and discuss the effect of the recent increase of interest rates on the future of inequality within our model.

The results of this chapter have been published in [69].

6.1 Posing the problem

Most studies consent (e.g. [60, 112, 65, 19, 38, 37, 48]) that in industrial countries the distribution of wealth reveals a two-tailed structure: Whereas for the majority of the population (95-99%) the empirical distribution can be well-described by a light-tailed or log-normal distribution, a power-law distribution turns out to be more suitable for the richest within an economy. In fact, the Pareto-distribution was initially suggested by Vilfredo Pareto [106] in 1897 to describe the distribution of income and wealth. Moreover, wealth is distributed significantly more unequal than income (see [34] or [112] and references therein). Figure 1.1 shows the distribution of net personal wealth for Germany and the USA, which confirms the two-tailed structure since the richest seem to follow a power-law distribution. Least square fit estimates a Pareto exponent of approximately 1.44 for Germany 2021 with similar values for 2011 and the USA. This exponent is similar for other countries as presented in [125] and widely stable in time [19], as already predicted by V. Pareto himself. For comparison, [38] also finds a Pareto-tail in income distribution with exponent varying between 2.42 and 3.96 in Germany since 1990, underlining the mentioned gap between income and wealth distribution.

Numerous models have been proposed to find the determinants of the power-law tail, some of which can be found in [60, 65, 19, 38, 29, 129, 85, 20, 35, 90, 130, 25, 99] and references therein. The early model proposed in [120], which was recently taken up in [76], already contains the idea of combining independent and reinforced elements. An essential reason for the power-law tail has been found in so-called increasing returns, i.e. the return rates on capital depend on the amount of capital a person or household owns. Detailed theoretical and empirical information on this phenomenon can be found e.g. in [5, 53, 60, 7, 51].

An established model for growth processes subjected to increasing returns is the non-linear Pólya urn model from the previous chapters, as promoted by Arthur [5]. [124] fits the classical generalized Pólya urn to American data. Indeed, this model creates Pareto-tailed wealth distributions (see Chapter 5), but a major drawback of this model is the occurrence of strong monopoly, i.e. from some point on only one (random) agent wins in all following steps. The aim of this chapter is to extend the generalized Pólya urn model, such that the empirical wealth distribution from Figure 1.1 emerges as a stable long-term distribution under the dynamics of our new model.

First, we pick up the main idea of Pólyas urn model and distribute the wealth created in an economy step by step among a fixed number $A \in \{2, 3, \ldots\}$ of agents. But in this chapter, we distinguish two different mechanisms of assigning an abstract unit of additional wealth to an agent. Assume that a company generates a unit of additional wealth, which corresponds to the gross yield. A certain share $r \in [0, 1]$ (the so-called "labor share") of the gross yield is payed to the employees via wages. The remaining $1 - r$ units (the "profit share") are assigned to the shareholders, either by paying dividends or by increasing the fundamental value of the company. For simplicity, we assume wages to be fixed in time and independent of wealth in our model, i.e. this share of the abstract wealth unit is distributed among all agents proportionate to some fixed vector. The other part of the added wealth unit is distributed among the agents via capital returns and does hence depend on the current wealth of the agents. Capital returns will be modeled as a non-linear Pólya urn as in Section 2.1, which includes the phenomenon of increasing returns. Hence, the remaining $(1 - r)$-share of the wealth unit is fully assigned to one randomly chosen agent. The share r is a measure for the importance of capital for the accumulation of wealth, but it differs from the official labour share arising in national accounting. We will discuss its precise meaning in Subsection 6.4.2. In more casual words: r regulates in how far it is possible to become rich through hard work. As explained in Subsection 6.4.2, it is justifiable to assume that the labor share is constant in time. Since the distribution of wages is an exogenous parameter in our model, we will particularly focus on providing an explanation for the discrepancy between wage distribution and wealth distribution as a consequence of increasing returns.

6.2 The model and some rigorous results

For the purposes of this chapter, we extend the model from Section 2.1 as follows. Again, let $A \in \{2, 3, \ldots\}$ be the number of agents in our economy and $r \in [0, 1]$ denotes the fraction of the added wealth, which is distributed proportionate to a deterministic vector $\gamma = (\gamma_1, \ldots, \gamma_A) \in \Delta_{A-1}$, representing the effect of wages on the accumulation of wealth. We fix a feedback function $F_i \colon (0, \infty) \to (0, \infty)$ for each agent $i \in [A]$. Then we adapt the definitions from Section 2.1 by defining a homogeneous Markov process $X(n) = (X_1(n), \ldots, X_A(n))$, $n \in \mathbb{N}_0$ in the state space $[1, \infty)^A$ with initial condition $X(0) \in [1, \infty)^A$ and transition probabilities given by

$$\mathbb{P}\left(X(n+1) - X(n) = v^{(i)} \,\middle|\, X(n)\right) = p_i\left(N + n, \frac{X(n)}{N + n}\right) := \frac{F_i(X_i(n))}{F_1(X_1(n)) + \ldots + F_A(X_A(n))} \tag{6.1}$$

for $i \in [A]$, where $v^{(i)} := (1 - r)e^{(i)} + r\gamma$. As in Section 2.1, $N := X_1(0) + \ldots + X_A(0)$ denotes the total wealth at time $n = 0$ and the corresponding process of wealth shares is defined as

$$\chi(n) := (\chi_1(n), \ldots, \chi_A(n)) := \frac{1}{N + n} X(n) \in \Delta_{A-1}, \quad n \in \mathbb{N}_0, .$$

For $r = 0$, this process coincides with the generalized Pólya urn from Section 2.1, whereas for $r = 1$ the process $\chi(n)$ is deterministic. For general $r \in (0, 1)$, this process can be considered as an extension of a generalized Friedman urn (see e.g. [107] an references therein) to non-linear feedback, which has not been studied before to our knowledge. In a generalized Friedman urn, a ball is drawn uniformly from an urn and, depending on its colour, the ball is replaced by a set of other balls, which may contain several different colours. Our process $(X(n))_n$ picks up this idea and includes additionally a non-proportionate drawing rule.

Corresponding to (4.1), define the vector field

$$G(N + n, x) := \mathbb{E}\left[X(n+1) - X(n) \,\middle|\, X(n) = \lfloor (N + n)x \rfloor\right] - x = (1 - r)p(N + n, x) + r\gamma - x \tag{6.2}$$

of centered expected increments for $x \in \Delta_{A-1}, n \in \mathbb{N}$ and $p(n, x) := (p_1(n, x), \ldots, p_A(n, x))$. In particular, the field G represents the expected increment of shares up to scaling, i.e.

$$\mathbb{E}\left[\chi(n+1) - \chi(n) \,|\, \chi(n) = x\right] = \frac{G(N + n, x)}{N + n + 1} \,. \tag{6.3}$$

Note that the time-scale of our model is non-linear, i.e. one step of the process does not correspond to a fixed period of time in reality. When $\mu = \mu(t)$ is the annual growth rate of our economy (given as an exogenous parameter), then we could instead consider the time-changed process $t \mapsto X\left(\lfloor \left((1 + \mu)^t - 1\right) N \rfloor\right)$, where t is time measured in years. This will be discussed in detail in Section 6.4, in the following the explicit time scale is not important.

A rigorous approach to the long time behaviour of this process is provided by the method of stochastic approximation, which we already used in Chapter 4. For that, we consider Doob's decomposition

$$\chi(n) = \chi(0) + H(n) + M(n) \,, \tag{6.4}$$

where

$$H(n) := \sum_{k=0}^{n-1} \frac{G(N+k, \chi(k))}{N+k+1} \quad \text{and} \quad M(n) := \sum_{k=0}^{n-1} \frac{1}{N+k+1} \xi(k),$$

with $\xi(n) := X(n+1) - X(n) - G(N+n, \chi(n)) - \chi(n)$. The process $H(n)$ is predictable with respect to the filtration $(\mathcal{F}_n)_n$ generated by the process $\chi(n)$. Moreover, the $\xi(n)$ are centered, bounded and uncorrelated since

$$\mathbb{E}\left[\xi_i(n)\xi_j(m)\right] = \mathbb{E}\left[\xi_i(n)\mathbb{E}\left[\xi_j(m) \,\middle|\, \mathcal{F}_n\right]\right] = 0 \quad \text{for } m > n,\, i, j \in [A]\,.$$

Hence, $M(n)$ is a martingale, which is bounded in L^2 and consequently almost surely convergent for $n \to \infty$.

For simplicity, we will mainly consider homogeneous feedback functions in the following, i.e. $F_i(k) = \alpha_i k^\beta$ for some $\alpha = (\alpha_1, \ldots, \alpha_A) \in (0, \infty)^A$ and $\beta \in \mathbb{R}$. This kind of feedback is particularly simple since the transition probabilities $p(n, x)$ do not depend on n, such that we can again establish the notation

$$p(x) := p(n, x) \quad \text{and} \quad G(x) := G(n, x) \tag{6.5}$$

for the homogeneous case with

$$G_i(x) = (1-r)\frac{\alpha_i x_i^\beta}{\alpha_1 x_1^\beta + \ldots + \alpha_A x_A^\beta} + r\gamma_i - x_i\,.$$

Exploiting the convergence of the martingale $M(n)$, one can show that the long time limits of this process are given by the zeros x of the vector field G, i.e. $G(x) = 0$, which we will refer to as **fixed points** of the dynamics.

Theorem 6.1. *For all $i \in [A]$ let $F_i(k) = \alpha_i k^\beta$ for $\alpha_i > 0, \beta \in \mathbb{R}$, $r \in [0,1]$. Then $\chi(n) \to \chi(\infty)$ converges almost surely to a stable fixed point of G for $n \to \infty$.*

Proof. The proof follows similar stochastic approximation arguments like in [107, 18, 30, 100]. Define the set $S \subset \Delta_{A-1}$ of fixed points of G. Similar to [18], a Lyapunov function for G is given by

$$L(x) := -(1-r)\log\left(\sum_{i=1}^A \alpha_i x_i^\beta\right) - r\sum_{i=1}^A \gamma_i \log x_i + \sum_{i=1}^A x_i \quad \text{for } x = (x_1, \ldots, x_A) \in \Delta_{A-1}^o \tag{6.6}$$

as $\frac{d}{dx_i}L(x) = -\frac{1}{x_i}G_i(x)$ and consequently $\langle \nabla L(x), G(x)\rangle = -\sum_{i=1}^A x_i \left(\frac{d}{dx_i}L(x)\right)^2 \leq 0$ with equality if and only if $x \in S$. Now, we observe that $L(\chi(n))$ eventually becomes a supermartingale:

$$\begin{aligned}
\mathbb{E}\left[L(\chi(n+1) - L(\chi(n)) \,\middle|\, \chi(n)\right] &= \mathbb{E}\left[\nabla L(\chi(n))\left((\chi(n+1) - \chi(n)) + O(1/n^2)\,\middle|\, \chi(n)\right]\right. \\
&= \nabla L(\chi(n))\mathbb{E}\left[\chi(n+1) - \chi(n) \,\middle|\, \chi(n)\right] + O(1/n^2) \\
&= \frac{1}{N+n+1}\langle \nabla L(\chi(n)),\, G(\chi(n))\rangle + O(1/n^2)
\end{aligned}$$

Since L is bounded from below, we get almost sure convergence of $L(\chi(n))$ from the martingale convergence theorem. Take an open ϵ-neighborhood $U_\epsilon \subset \Delta_{A-1}$ of S. Then there is $\delta(\epsilon) > 0$ such that

$$\mathbb{E}\left[L(\chi(n+1)) - L(\chi(n)) \,\middle|\, \chi(n) = x\right] < -\frac{\delta(\epsilon)}{n} \quad \text{for all } x \in \Delta_{A-1} \setminus U_\epsilon \,,$$

if n is large enough. Due to the divergence of the harmonic series, the limit point needs to be in any U_ϵ and consequently in S.

The non-convergence to unstable fixed points is technically more demanding and follows basically from arguments like in [30, Lemma 5.2.], [108, Theorem 1] or [100, Chapter 5]. Note that the stable fixed points of G are just the strict local minima of L. Maxima and saddle points of L, i.e. unstable fixed points of G, are not attained as limit points of $L(\chi(n))$ due to noise of order $\frac{1}{n}$. □

Like in Theorem 2.5, it is possible to extend Theorem 6.1 to inhomogeneous feedback functions, provided that the field $G(k, x)$ converges for $k \to \infty$ sufficiently fast. For our applied purposes, these inhomogeneous feedback functions do not grant any enriching insights, such that we will neglect them.

Also note that the equation $G(x) = 0$ does in general consist of $A - 1$ equations for $A - 1$ variables, since the A-th equation is redundant due to

$$\sum_{i=1}^{A} G_i(x) = 0 \quad \text{and} \quad \sum_{i=1}^{A} x_i = 1 \quad \text{since } x \in \Delta_{A-1} \,.$$

Hence, heuristically, the zero-set of G can be considered to be discrete.

An interesting observation in the situation of Theorem 6.1 is that the limiting share of all agents $i \in [A]$ with positive wage $\gamma_i > 0$ is bigger than $r\gamma_i$, i.e. $\chi(\infty) > r\gamma_i$ almost surely for all $i \in [A]$ since $p(n, x) > 0$ for all $x \in \Delta^o_{A-1}$. The inequality is strict since all agents do not only receive their wage, but also capital returns on their savings.

As mentioned before, our process is deterministic for $r = 1$. The following proposition states that the process reveals a deterministic long-time behaviour even for large enough $r < 1$. In that case, agents with zero wage will have vanishing shares on the long run.

Proposition 6.2. *For all $i \in [A]$ let $F_i(k) = \alpha_i k^\beta$ for $\alpha_i > 0$, $\beta \geq 1$. Then there is a **critical labor share** $r_c < 1$, such that for all $r \geq r_c$ the limit $\chi(\infty) := \lim_{n\to\infty} \chi(n)$ is deterministic. If $r \geq r_c$, then any agent $i \in [A]$ with $\gamma_i = 0$ necessarily fulfills $\chi_i(\infty) = 0$.*

Proof. First, Theorem 6.1 implies the existence of a fixed point of G. For $r \in (0, 1)$ take $x_r, y_r \in \Delta^o_{A-1}$ satisfying $G(x_r) = G(y_r) = 0$. Define $G_0(x) := p(x) - x$ for $x \in \Delta_{A-1}$. Note that the field G_0 does not depend on r. Then:

$$\begin{aligned} G_0(x_r) - G_0(y_r) &= \frac{1}{1-r}\left((1-r)p(x_r) + r\gamma - x_r - (1-r)p(y_r) - r\gamma + y_r + rx_r - ry_r\right) \\ &= \frac{1}{1-r}\left(G(x_r) - G(y_r) + rx_r - ry_r\right) = \frac{r}{1-r}(x_r - y_r) \end{aligned}$$

As $\beta \geq 1$, G_0 is Lipschitz-continuous with a Lipschitz-constant $L = L(\alpha, \beta) < \infty$. Hence, we have $x_r = y_r$, when $\frac{r}{1-r} > L$, and as a consequence G has only one unique fixed point for $r \geq r_c := \frac{L}{1+L} \in (0, 1)$.

Now, let $r \geq r_c$ and assume (w.l.o.g.) $\gamma_1 = 0$. If $y_0 \in \Delta_{A-2}$ is a zero of the restricted field $\tilde{G}(y) = G(0, y)$, $y \in \Delta_{A-2}$, which corresponds to a system with $A-1$ agents, then $(0, y_0)$ is a zero of G. Uniqueness of the fixed point for $r \geq r_c$ and the existence of such y_0 imply $\chi_1(\infty) = 0$. □

This proof also implies that there is no further (unstable) fixed point of G for $r \geq \frac{L}{1+L}$, but there might be unstable fixed points for general $r \geq r_c$ as visible in Figure 6.3. Note that Proposition 6.2 may hold with $r_c < \frac{L}{L+1}$, our argument provides only an upper bound for r_c.

For $r = 0$, our process equals the generalized Pólya urn from Section 2.1. If in addition $\beta > 1$, then the process reveals strong monopoly with probability one, i.e. at some point one agent wins all following steps. Of course, this cannot happen for $r > 0$ since all agents get at least their wage and receive capital returns on their wage on top. Nevertheless, we suggest to call agents **winner**, if their wealth share exceeds their income share, i.e. $\chi_i(\infty) > \gamma_i$. Otherwise we call them loser. In that sense, we can still identify a unique random winner for small enough $r > 0$.

Proposition 6.3. *Let $F_i(k) = k^\beta$ for all $i \in [A]$ with $\beta > 1$. Then there is $r'_c > 0$ such that for all $r < r'_c$*

$$\mathbb{P}\left(\exists!\, i \in [A] \colon \chi_i(\infty) > \gamma_i\right) = 1$$

and for all $i \in [A]$

$$\mathbb{P}\left(\chi_i(\infty) > \gamma_i\right) > 0\,.$$

Proof. Let $i \in [A]$. Denote by $DG(x) := \left(\frac{\partial G(x)}{\partial x_i \partial x_j}\right)_{i,j}$ the differential matrix of G in $x \in \Delta_{A-1}$. For $r = 0$, a simple computation shows $\nabla G_j(e^{(i)}) = (-\delta_{l,j})_{l=1,\dots,A}$ for all $j \in [A]$. Hence, $DG(e^{(i)})$ is negative definite and invertable. Then we get from the implicit function theorem that there is $\epsilon > 0$, $r'_c > 0$ such that for all $r < r'_c$ there is exactly one zero of G in the ϵ-neighborhood of $e^{(i)}$. Denote this fixed point by $x^{(i)}(r)$. Obviously, for all agents $j \neq i$ with $\gamma_j = 0$ we must have $x_j^{(i)}(r) = 0$ due to the uniqueness of the fixed point. Hence, assume without loss of generality that $\gamma_j > 0$ for all $j \neq i$ and suppose $\epsilon \leq \min\{\gamma_j \colon j \neq i\}$. Consequently, $x_j^{(i)}(r) \leq \gamma_j$ for all $j \neq i$ and $r < r'_c$.

It remains to show that for $r < r'_c$ there are no other stable fixed points of G. For that, consider the "r=0"-field $G_0(x) = p(x) - x$. We know from Section 3.1 that all zeros of G_0 have the form $x^{(S)} = \left(\frac{1}{\#S}\mathbb{1}_S(i)\right)_{i \in [A]}$ for a non-empty subset $S \subset [A]$. Since G is a continuous perturbation of G_0, we know that for small enough r all zeros of G are located in an ϵ-neighborhood of these points $x^{(S)}$. Moreover, $x^{(S)}$ is unstable for $\#S > 1$ such that $DG_0(x^{(S)})$ has at least one positive eigenvalue. Since eigenvalues of $DG(x)$ do continuously depend on x and r, there is still at least one positive eigenvalue of $DG(x)$ for any zero x of G that is located in an ϵ-neighborhood of $x^{(S)}$ with r small enough. Hence, all stable fixed points are close to an $x^{(S)}$ with $\#S = 1$ if r is small. The stability of these fixed points can be shown similarly using negative definiteness of $DG_0(e^{(i)})$. □

An interesting implication of the construction of our stable fixed points is the following: For $r < r_c'$, there is exactly one fixed point close to each corner of Δ_{A-1}. Hence, $\chi(\infty)$ is fully determined by picking the winner, i.e. there is no random hierarchy between the losers. In the next section, we will see that the fixed points disappear one by one, when r is increased, until finally only one fixed point remains for $r \geq r_c$.

Since it will be of particular interest in Section 6.5, let us now discuss the **linear case** $F_i(k) = \alpha_i k$ with skill parameter $\alpha_i > 0$, which corresponds to wealth independent return rates. Obviously, γ is the unique stable fixed point of G (6.5) when $\alpha_1 = \ldots = \alpha_A$ and $r > 0$ and hence, $\chi(n)$ converges to γ almost surely. In case of the standard Pólya urn (i.e. $\alpha_i = 1$ and $r = 0$), $\chi(n)$ converges almost surely towards a random point. For unequal α_i and $r = 0$, the process reveals a deterministic weak monopoly, i.e. $\chi(n)$ converges to $e^{(i)}$, where i is the agent with the largest α (see Proposition 3.30). For $r > 0$, there is a deterministic limit for all choices of $\alpha_i > 0$, in particular for equal α_i, as shown in the following proposition (see also Figure 6.3 (c)). This does also hold in the sub-linear case, where $\chi(\infty)$ is deterministic even for $r = 0$.

Proposition 6.4. *Let $r > 0$ and $F_i(k) = \alpha_i k^\beta$ for $\alpha_i > 0$ and $\beta \leq 1$. Then $\chi(n)$ converges almost surely to a deterministic point $\chi(\infty)$ for $n \to \infty$, i.e. $r_c = 0$.*

Proof. Using the argument from the proof of Theorem 6.1, we have to show that the Lyapunov function L defined in (6.6) has a unique minimum. For that, we prove that L is strictly convex. Direct calculation yields that the Hessian of L is of the form

$$\left(\frac{\partial L(x)}{\partial x_i \partial x_j}\right)_{i,j} = c(x) \cdot v \cdot v^T + A(x), \quad x = (x_1, \ldots, x_A) \in \Delta^o_{A-1}\,,$$

where $c(x) := (1-r)\beta^2 \left(\sum_{i=1}^A \alpha_i x_i^\beta\right)^{-2} \geq 0$ and $v = \left(\alpha_i x_i^{\beta-1}\right)_{i\in[A]} \in (0,\infty)^A$ and $A(x) = (A_{i,j}(x))_{i,j\in[A]}$ is a diagonal matrix with

$$A_{i,i}(x) = r\gamma_i x_i^{-2} + (1-r)\beta(1-\beta)\alpha_i^2 x_i^{\beta-2} \left(\sum_{j=1}^A \alpha_j x_j^\beta\right)^{-1} \geq 0$$

Assume $r < 1$ as $r = 1$ is trivial. Since $v \cdot v^T$ is non-negative definite, the Hessian of L is positive definite if either $\gamma_i > 0$ for all $i \in [A]$ or $\beta < 1$. But if $\beta = 1$, we can w.l.o.g. assume $\gamma_i > 0$ due to Lemma 6.5. Hence, L is strictly convex. □

Lemma 6.5. *Let $r > 0$ and $F_i(k) = \alpha_i k$ for $\alpha_i > 0$. Then $\chi_i(\infty) = 0$ for any agent $i \in [A]$ with $\gamma_i = 0$.*

Proof. Due to the linearity, if suffices to consider a system with $A = 2$ and $\gamma_1 = 1$, since this process is equivalent to the group-process $\left(\sum_{\substack{i\in[A]\\ \gamma_i>0}} \chi_i(n), \sum_{\substack{i\in[A]\\ \gamma_i=0}} \chi_i(n)\right)_n$. But then a simple calculation shows that

$$G(x) = 0 \quad \Leftrightarrow \quad (1-r)\frac{\alpha_1 x_1}{\alpha_1 x_1 + \alpha_2(1-x_1)} + r = x_1$$

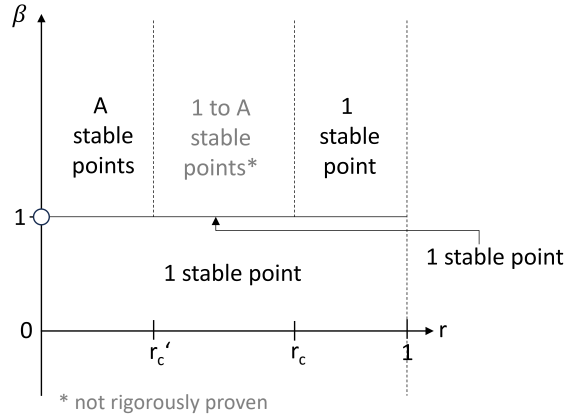


Figure 6.1: Qualitative illustration of the number of stable fixed points of G for homogeneous feedback and different r and β. ∘ marks the classical Pólya urn, which exhibits either deterministic (weak) monopoly or a Dirichlet distributed limit. The process has a deterministic limit point if and only if G has only one stable fixed point.

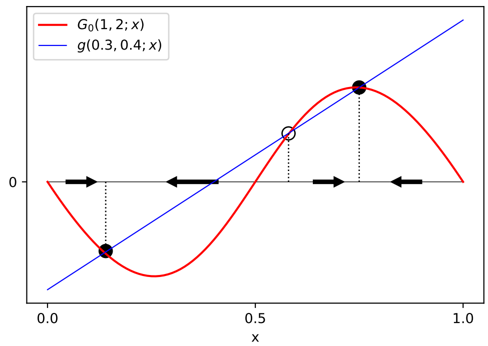
has only the solution $x_1 = 1$. □

Let us finally sum up the homogeneous case $F_i(k) = \alpha_i k^\beta$: The limit $\lim_{n\to\infty}\chi(n)$ does exist and has a discrete distribution (if not $\beta = 1$ and $\alpha_i = \alpha_j$), but both deterministic and random limits are possible depending on r and β. Remarkably, the set of possible limit points does not depend on the initial configuration $X(0)$, but the probability that a specific limit point is attained might depend on $X(0)$. This means for $r < r_c$ that agents with high wage and low initial wealth may never fully catch up their initial disadvantage, whereas they will do so for $r \geq r_c$. For $\beta > 1$ and $r < r_c$, the first steps of the process decide which limit point is attained, because the process behaves almost deterministic for large market sizes according to the LLN presented in Appendix C.1. For small enough r, we can still identify unique winners and any agent can be the winner. For constant or decreasing return rates $\beta \leq 1$, $\lim_{n\to\infty}\chi(n)$ is deterministic for all $r > 0$ and hence $r_c = 0$, whereas $r_c > 0$ in the increasing return case $\beta > 1$. Figure 6.1 finally illustrates these findings.

6.3 The two and three agent case

In order to gain a visual understanding of the long-time behaviour of this process, we will discuss the homogeneous case with $A = 2$ in detail. So, let $F_1(k) = k^\beta$, $F_2(k) = \alpha k^\beta$ for $\beta \in \mathbb{R}$ and $\alpha \geq 1$. For simplicity, we establish the notation $G(x) = G_1(x, 1-x)$, $\gamma = \gamma_1$ and $p(x) = p_1(x, 1-x)$ for $x \in [0,1]$, i.e. x represents the share of agent 1 and agent 2 has share $1-x$. Then

$$G(x) = (1-r)(p(x) - x) + r(\gamma - x) = 0 \quad \Leftrightarrow \quad p(x) - x = \frac{r}{1-r}(x - \gamma)$$



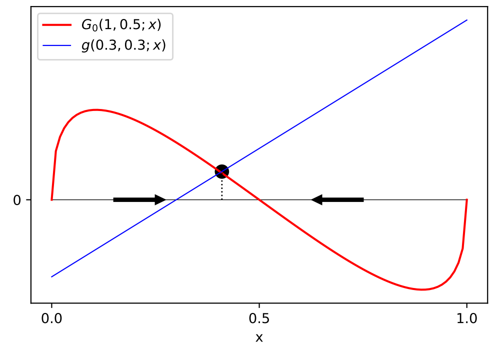


Figure 6.2: The line $g(r, \gamma; x)$ (6.8) and the field $G_0(\alpha, \beta; x)$ (6.7) against wealth x of agent 1. $\bullet$ marks stable and $\circ$ unstable fixed points. The arrows indicate the direction of the field G.

and the stable fixed points of G are the downcrossings of the "$r = 0$"-field

$$x \mapsto G_0(\alpha, \beta; x) := p(x) - x = \frac{x^\beta}{x^\beta + \alpha(1-x)^\beta} - x \tag{6.7}$$

with the line

$$x \mapsto g(r, \gamma; x) := \frac{r}{1-r}(x - \gamma) \ . \tag{6.8}$$

These downcrossings constitute the possible long-time limits of the process $\chi(n)$ according to Theorem 6.1. The upcrossings are unstable fixed points and are not attained as long-time limits.

The situation is qualitatively illustrated in Figure 6.2 and 6.3. An increase of the labor share r implies a larger slope of the line g, where the slope diverges for $r \to 1$. Changes of the relative wage γ result in a parallel shift of g. The impact of the fitness parameter α and the feedback strength β is included in the field G_0. Let us now have a closer look on the possible cases.

1. Figure 6.3 (a) shows the symmetric case $\alpha = 1, \gamma = 0.5$ with $\beta > 1$ for different labor shares r. It is apparent that $x = \frac{1}{2}$ is the only stable fixed point of G if and only if

$$\frac{d}{dx}(p(x) - x) = \beta - 1 \leq \frac{r}{1-r} \quad \Leftrightarrow \quad r \geq r_c := \frac{\beta - 1}{\beta} \ .$$

   For $r < r_c$, there are two stable fixed points which are symmetric w.r.t. $\frac{1}{2}$. For $r \to 0$, the two fixed points converge to 0 resp. 1, consistent with the strong monopoly for $r = 0$. The critical share r_c is increasing in β due to the stronger feedback and converges to 1 for $\beta \to \infty$.

   More explicit for $\beta = 2$, we have $r_c = \frac{1}{2}$ and for $r < \frac{1}{2}$ the two stable fixed points are given by $\frac{1}{2} \pm \sqrt{1 - 2r}$. For other choices of α, β explicit expressions are lengthy or not known.

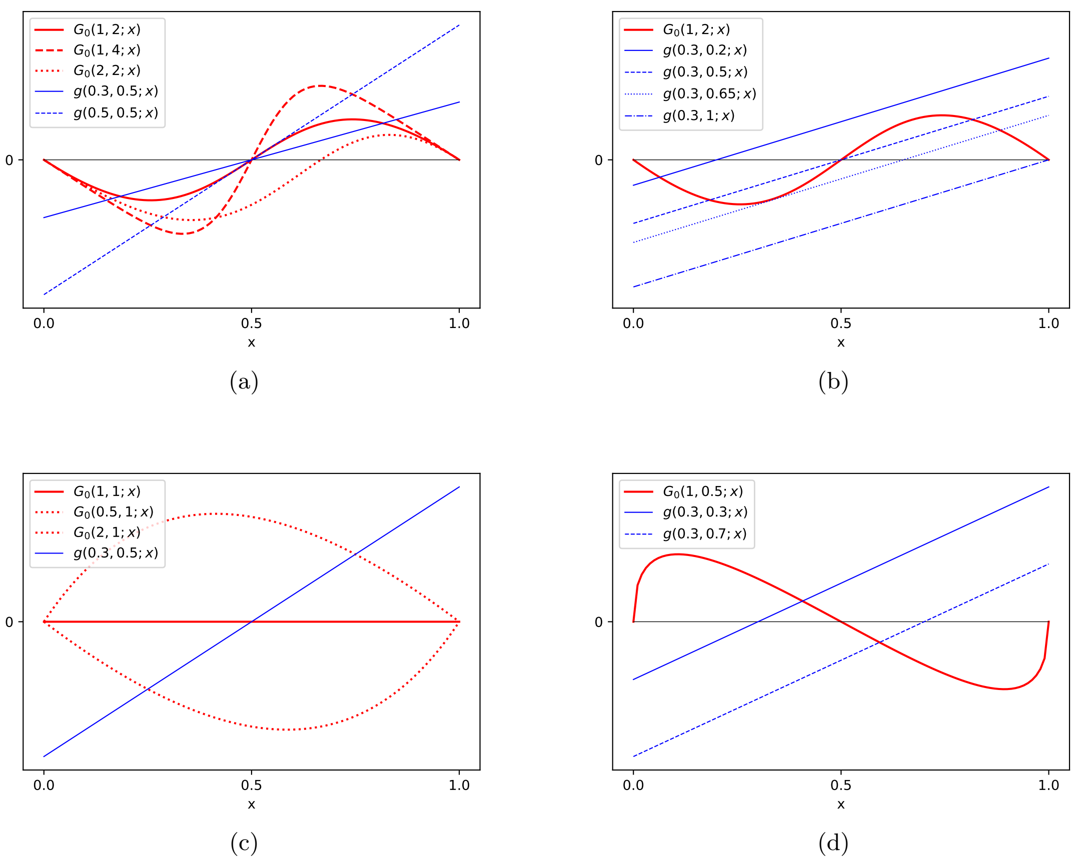


Figure 6.3: The line $g(r, \gamma; x)$ (6.8) and the field $G_0(\alpha, \beta; x)$ (6.7) against wealth x of agent 1 for various parameters.

In the asymmetric case $\alpha > 1$, where agent 2 is fitter than agent 1, we observe in general a shift of the stable fixed points towards agent 2. Moreover, the critical share r_c is smaller than for $\alpha = 1$.

2. Figure 6.3 (b) illustrates the situation with equal fitness $\alpha = 1$ of agents, varying wage distribution γ and fixed $r > 0$. First, we note that the critical rate r_c is smaller when wages are distributed unequally $\gamma \neq 0.5$, i.e. for fixed r we can obtain either random or deterministic limits depending on γ.

   Second, we observe that for $\beta > 1$ the long-time wealth is distributed more unequal than wages. To be more precise, if $\gamma < \frac{1}{2}$, then $\chi_1(\infty) < \gamma$ and vice versa. The gap between γ and $\chi_1(\infty)$ is bigger the smaller r and the larger β is.

   Third, for fixed $r < r_c$, there there two choices of γ, such that **saddle points** occur (see e.g. line $\gamma''' \approx 0.65$). These saddle points are stable from one side (from the left in Figure 6.3), but unstable from the other side. Hence, the process may stick to these points for a long time, but will finally escape towards the only fully stable point due to noise.

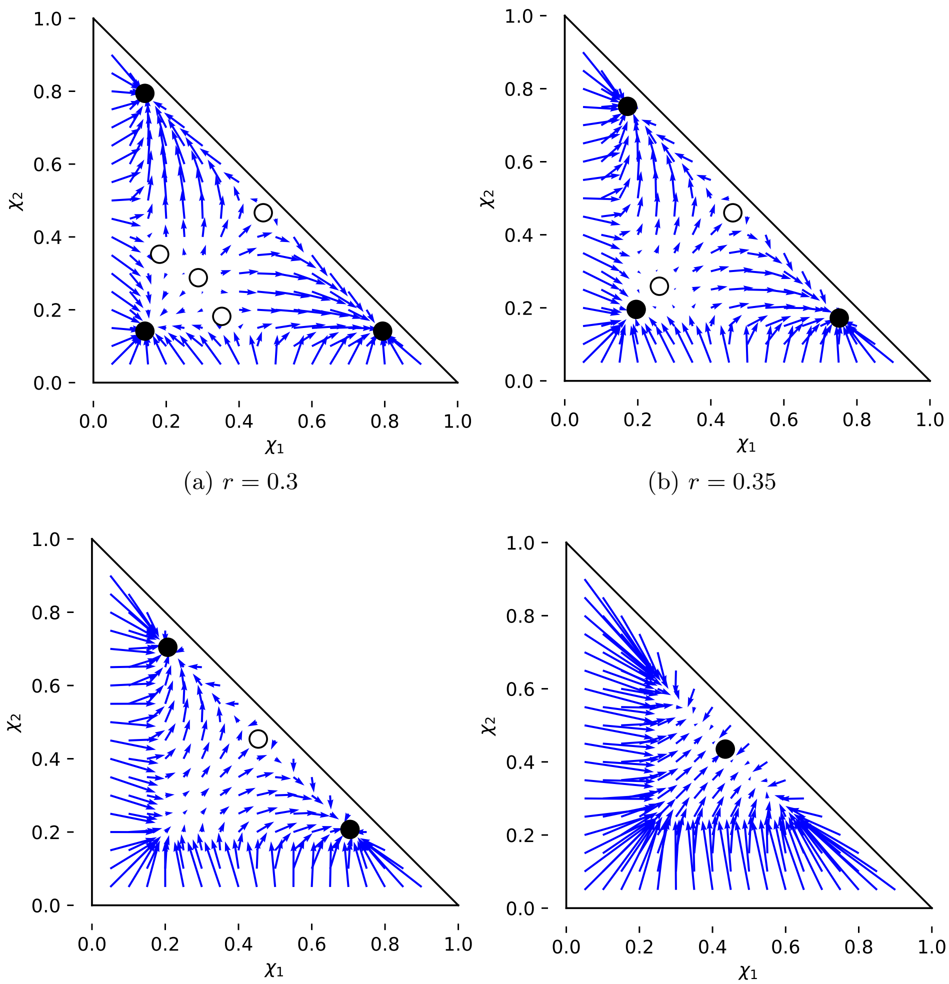
   Finally, for $\gamma = 1$, weak monopoly of agent 1, i.e. $\chi_1(\infty) = 1$, is possible with positive probability. But for $r < r_c$, both weak monopoly of agent 1 and positive shares for both agents are possible, depending on who wins the first steps of the process.

3. Figure 6.3 (c) shows the linear case $\beta = 1$, where we have unique fixed points whenever $r > 0$, such that $r_c = 0$. For $\alpha = 1$, the fixed point is simply γ. Recall that for $r = 0$ the limiting share $\chi_1(\infty)$ has a beta distribution. Changes of α for $r > 0$ result to a distortion of the unique fixed point towards the fitter agent.

4. Figure 6.3 (d) shows the situation for $\beta < 1, \alpha = 1$, where we still have $r_c = 0$. This situation corresponds to decreasing return rates on capital. Here, wealth is distributed more equally than wages, i.e. we have $\chi_1(\infty) > \gamma$ whenever $\gamma < \frac{1}{2}$ and $r > 0$.

Moreover, it becomes apparent from Figure 6.3 that we always have $r'_c = r_c$ when $A = 2$. In 2. we already mentioned the occurrence of saddle points, which the process may approach and remain there for long time, but will eventually leave. Although this behaviour seems to be rare in the two agent case as they do only occur for specific pairs of γ, r, these points are far more important in larger systems. Hence, we also have a close look on the $A = 3$ case in order to deepen our understanding of the long time behaviour of our process. Figure 6.4 shows the field G for asymmetrical wage vector and varying r. We return to the original notation of G introduced in Section 6.2.

For small r (Figure 6.4 (a)), we have 7 fixed points, where the three stable ones are close to the corners of the simplex, i.e. one random agent wins the bulk of the wealth. Casually speaking, we will call these **monopoly fixed points** in the following. Note that this concept does not necessarily coincide with the definition of "winners" above. Moreover, there are three saddle points, where basically two agents fairly share the total wealth. For appropriate starting points, the process may first approach these points, but finally converge to one of the



(a) $r = 0.3$ (b) $r = 0.35$

(c) $r = 0.4$ (d) $r = 0.55$

Figure 6.4: The field G (see (6.5)) with $A = 3$, $\beta = 2, \alpha_1 = \alpha_2 = \alpha_3 = 1$ and $\gamma = (0.4, 0.4, 0.2)$ for various labor shares r. Stable fixed points are marked with $\bullet$ and unstable fixed points with $\circ$. Their exact position has been computed explicitly with Mathematica.

stable points. In addition, there is one more fully repelling point in the middle. Note that this situation is similar to the $r = 0$ case discussed in the previous chapters.

When we now slowly increase r to some moderate level (Figure 6.4 (b)), we observe that at first the saddle points, where agent 3 (the one with the lowest wage) is involved, disappear, but the monopoly fixed point of agent 3 still exists. When r is increased even more (Figure 6.4 (c)), then the monopoly fixed point of agent 3 also disappears and only the monopoly fixed points of the agents with higher wage remain together with their saddle point. Finally, when r becomes larger than r_c (≈ 0.55 in the situation of Figure 6.4), then the two remaining stable fixed points merge and the process converges to a deterministic point.

Heuristically, we can generalize this visual grasp to **larger systems** $A > 3$ as follows. When r is small, then we have A stable fixed points, which are close to the corners of the simplex. Moreover, for any subset $S \subset \{1, \dots, A\}$, $\#S > 1$, there is a saddle point, where the bulk of the total wealth is shared between the agents in S. The saddle points are not attained as long-time limits of the process, but they may dominate the transient behaviour of the system which is relevant in practice. In total, there are $2^A - 1$ fixed points for small r. When r is increased, fixed points shift towards the middle of the simplex and disappear one by one, where the monopoly fixed points of agents with low wage as well as the saddle points referred to them disappear first. r_c is the minimal share such that only one of the fixed points survives and r'_c is the labor share where the first monopoly fixed point disappears. Consequently for "moderate" r, the process converges to a random monopoly fixed point, but only agents with large enough wage can be the monopolist. With this heuristic, we conjure that there are at most $2^A - 1$ fixed points in total and that at most A of them are stable. Appendix C.3 adds some more heuristics on the number and position of the zeros of G.

6.4 Simulations for homogeneous feedback

The goal of this section is to find a parameterization of the model introduced in Section 6.2, such that the real distribution of wealth in Germany 2021 is reproduced as well as possible by simulation. We denote by

$$CDF_{ger} : \mathbb{R} \to [0, 1] \quad \text{the empirical distribution function of wealth, Germany 2021} \tag{6.9}$$

according to the data from [111] (shown in Figure 1.1) and compare it to the simulated wealth distribution function CDF_{sim} defined in (6.10). Currently, about 70 million adults are living in Germany, but the data on wealth distribution from [111] have a much lower resolution. Simulating a system with millions of agents would therefore be computationally very demanding with essentially no verifyable benefit. Therefore, we aggregate and take $A = 10,000$, such that each agent represents 0.01% of the adults in reality. In particular, the k-th richest agent represents the $(k-1)/10,000$ to $k/10,000$ quantile in reality, which are about $7,000$ adults.

According to [111], the average net personal wealth per adult in Germany 2021 amounts to 227,567€. One unit of wealth in our model corresponds to 10€ in reality, which is a rather fine resolution. Note that our model is also scale-invariant in the sense that transition probabilities (6.1) are invariant under a rescaling of $X(n)$, therefore the choice of wealth units is not critical. We will simulate $n = 280,000,000$ steps of our process, such that the average wealth after n steps equals approximately the average wealth in reality. We will see in Figure 6.6 that the wealth distribution is fairly stable after n steps, so that distributing wealth in smaller units would would not yield any further insight. The wealth distribution after simulating the model (6.1) for n steps is then given by

$$CDF_{sim}\colon \mathbb{R} \to [0,1],\, w \mapsto \frac{1}{A}\sum_{i=1}^{A} \mathbb{1}_{\{10X_i(n)\leq w\}}\ . \tag{6.10}$$

We aim to reproduce the CDF_{ger} from generic small initial data and take an initial configuration $X(0)$, such that each agent has only one unit on average. Recall that the set of possible limiting wealth distributions is independent of $X(0)$ as explained in Section 6.2, but the probabilities that a certain limit point is attained does depend on $X(0)$.

In this section, we first consider symmetric and homogeneous feedback, i.e. $F_i(k) = k^\beta$ for some $\beta > 1$, and set $\alpha_i = 1$ for all agents, i.e. we neglect the effect of personal skills on capital returns. Of course, this is a simplifying assumption as there is probably a positive correlation between investment skills and wages which we will investigate in Section 6.5. To first approximation, it appears justified to assume that similarly affluent agents invest their money similarly and therefore achieve similar expected return rates. This is in particular plausible since each agent in our simulation represents 7,000 people in a corresponding wage-class in reality, so we can not account for completely untypical behaviour of some individuals anyway. In addition, we can think of people to improve their investment skills the more capital they have for investment, which are thus correlated with wealth (as captured by our parameter $\beta > 1$) rather than with wages. In Subsections 6.4.1, 6.4.2 and 6.4.3 we fit the parameters β, γ and r and finally show simulation results in Subsection 6.4.4.

6.4.1 The wage-vector γ

In our model, the distribution of wages is an exogenous parameter, which is invariant in time and is represented by the normalized vector γ. Wages have a significant impact on the wealth of poor agents, whereas the wealth of the rich is mainly determined by capital returns. As this chapter focuses on modelling the wealth distribution of the rich, i.e. the power law tail mentioned in Section 6.1, we are content with a rather rough wage-model. We will use data on German wages in 2018 as given in [31] (see Figure C.1 (a)). Note that only the shape of the wage-distribution is relevant since γ is normalized. The wage distribution is derived from data on income tax and contains income from employed and self-employed labor. Capital returns are basically not included as it is not subject to income tax in Germany (there is a flat-rate tax instead). The only exception worth mentioning is rental income, which only poses 2% of total income and can hence be neglected, too. Let $\tilde\gamma_1, \ldots \tilde\gamma_A$ be drawn independently from the distribution shown in Figure C.1 (a), where we assume uniform distributions within the intervals. For the agents with wage $> 1,000,000$ Euro, we suppose an exponential tail and thus all $\tilde\gamma_i$ are distinct.

In order to generate realistic wealth distributions in our model, it is important to distinguish that wages can either be used for consumption or for investment. Hence, we are less interested in the pure wage distribution modelled by $\tilde{\gamma}$ than in the distribution of savings, which add to the wealth of an agent and generate capital returns. It stands to reason that the agents with lowest wage need essentially all of it for consumption, whereas the highest-payed agents can invest almost their entire wages since even luxury consumption like jewelry and real estate increases value. For simplicity, we assume a linear relationship between the saving rate and the index of ordered wages $\tilde{\gamma}_{1:A} < \ldots < \tilde{\gamma}_{A:A}$, such that the agent with rank $i{:}A$ saves a fraction $\frac{i}{A}$ of their wage. Hence, we define the normalized sample γ as

$$\gamma_i := \frac{\tilde{\gamma}_{i:A}\, i/A}{\sum_{j=1}^{A} \tilde{\gamma}_{j:A}\, j/A} = \frac{i\tilde{\gamma}_{i:A}}{\sum_{j=1}^{A} j\tilde{\gamma}_{j:A}} \quad \text{for all } i \in [A] \ .$$

More detailed information on saving rates depending on income in Germany can be found in e.g. [22], confirming our linear interpolation as an appropriate approximation. Of course, there is much more refined research on the distribution of income like [38], most of which include capital returns in their data and are therefore not suitable for our purpose.

6.4.2 The labor share r

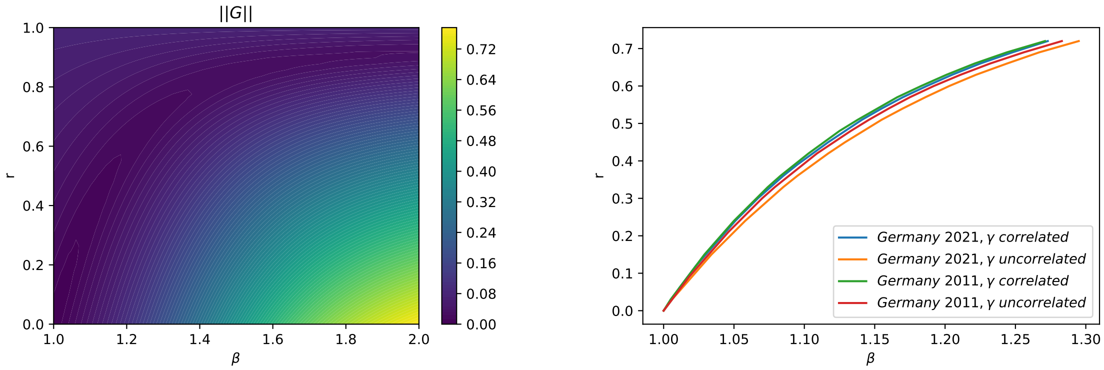
In official macroeconomic accounting, the labor share is defined as the part of the national income allocated to wages, which fluctuates in Germany between 64% and 72% since reunification 1991 [33]. In our model, however, the parameter r rather represents the part of the wealth increase that is due to savings from wages. Hence, it is not useful to simply set $r \approx 0.7$ for several reasons. First, national income does not encompass an increasing value of existing assets like real estate or corporate stocks, which reinforces the significance of capital on the growth of personal wealth. Second, the national income contains consumption, which does not add to wealth. The share used for consumption is presumably higher for wages than for capital returns, which again increases the significance of capital returns for wealth aggregation. Third, the effect of different taxation is not taken into account in the official labor share.

As a consequence, we estimate the parameter as

$$r = \frac{\text{average net wage} * \text{average savings rate}}{\text{increase of average wealth}} \tag{6.11}$$

for a fixed period of time. For the increase of average personal wealth, we take data from [111] again. [32] provides information about average net wages and saving rates.

Figure C.1 (b) shows empirical values of r according to formula (6.11) for several years. We observe extreme peaks in 2009 and 2020, which are due to the financial resp. the Covid crisis, where the increase of wealth was small, whereas wages are less sensitive to such events. Before 2020, the saving rate fluctuated slightly around 10%. Between 2013 and 2019, the empirical r values are stable between 20% and 27% percent. This low level is due to strong increases in value of real estate and stocks, caused by zero interest politics. Before the financial crisis, our adjusted labor share widely coincided with the official share. In the following, we will mostly use $r = 0.3$, but also consider higher values and show in detail how they affect our results in Section 6.6.



(a) (b)

Figure 6.5: The left figure shows a contour plot of $\|G(x_{ger})\|$ with a normalized sample x_{ger} from CDF_{ger} (6.9) for different values of r and β, where γ and x_{ger} are fully correlated. The right one shows the minimizing $r-\beta$-line for different years where γ and x_{ger} are fully correlated or uncorrelated.

6.4.3 The reinforcement parameter β

After fixing the parameters γ and r, we finally have to find an appropriate choice for β. The parameter β regulates the reinforcement mechanism of increasing returns, i.e. $\beta = 1$ corresponds to constant expected return rates and $\beta > 1$ corresponds to increasing returns. Hence, reinforcement for $\beta > 1$ determines the deviation of the wealth distribution from the wage distribution. We will estimate β directly from the shape of the desired wealth distribution shown in Figure 1.1 by adjusting β such that CDF_{ger} is fairly stable under the dynamics (6.1), since the empirical wealth distribution can be considered as stable in time up to scaling. From (6.2), it is easy to see that for any fixed β and x, there is a unique r minimizing $\|G(x)\|$.

Proposition 6.6. *For any fixed $\beta \in \mathbb{R}, x, \gamma \in \Delta_{A-1}$, the Euclidean norm $\|G(x)\|$ is minimal when*

$$r = r^\star(\beta) := 1 - \frac{\langle p(x) - \gamma, x - \gamma \rangle}{\|p(x) - \gamma\|^2}$$

If $r^\star(\beta) < 0$ (resp. > 1), it has to be replaced by 0 (resp. 1).

Proof. Define $w = p(x) - x$ and $v = \gamma - x$, such that $G(x) = (1-r)w + rv$ is a convex combination of v and w. Then:

$$\begin{aligned}\frac{d}{dr}\|G(x)\|^2 &= \frac{d}{dr}\sum_{i=1}^{A}((1-r)w_i + rv_i)^2 = 2\sum_{i=1}^{A}((1-r)w_i + rv_i)(v_i - w_i)\\ &= 2\sum_{i=1}^{A}\left(r(v_i - w_i) + w_i\right)(v_i - w_i) = 2r\|v-w\|^2 + 2\langle w, v-w\rangle\\ &= 2r\|\gamma - p(x)\|^2 + 2\langle p(x) - x, \gamma - p(x)\rangle\end{aligned}$$

Since $\|G(x)\|^2$ is a non-negative quadratic polynomial in r, the unique minimum is

$$r = \frac{\langle p(x) - x, p(x) - \gamma\rangle}{\|p(x) - \gamma\|^2} = 1 - \frac{\langle x - \gamma, p(x) - \gamma\rangle}{\|p(x) - \gamma\|^2}\,.$$

□

$r^\star(\beta)$ can be interpreted as the orthogonal projection of $x-\gamma$ on $p(x)-\gamma$. When $x-\gamma$ and $p(x)-\gamma$ are negatively correlated, i.e. $\langle x-\gamma, p(x)-\gamma\rangle < 0$, then $r=1$ is optimal. If $\|p(x)-\gamma\| < \|x-\gamma\|$ and the angle between $x-\gamma$ and $p(x)-\gamma$ is small, then $r=0$ is optimal. Moreover, if $(x_i,\gamma_i)_{i\in[A]}$ is a normalized sample of a positive random vector (X,Γ), then

$$r^\star(\beta) \approx 1 - \frac{Cov\left(\frac{X}{\mathbb{E}X} - \frac{\Gamma}{\mathbb{E}\Gamma}, \frac{X^\beta}{\mathbb{E}X^\beta} - \frac{\Gamma}{\mathbb{E}\Gamma}\right)}{\mathbb{E}\left(\frac{X^\beta}{\mathbb{E}X^\beta} - \frac{\Gamma}{\mathbb{E}\Gamma}\right)^2}$$

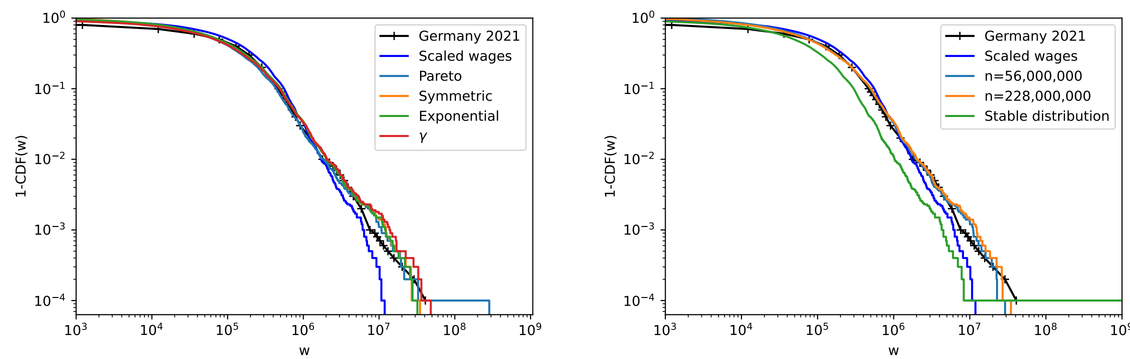
for large A. Hence, $r^\star(\beta)$ is asymptotically (for $A\to\infty$) independent of A.

Figure 6.5 (a) shows a contour plot of $\|G(x_{ger})\|$ for different choices of r,β, where $x_{ger}\in\Delta_{A-1}$ is a normalized sample from the empirical wealth distribution CDF_{ger} (6.9). This indicates that the relation between the parameters is indeed one-by-one, i.e. for any given r there is exactly one optimal β. Figure 6.5 (b) underlines, that the resulting r-β-line is fairly stable in time and not very sensitive on the correlation between wage γ_i and wealth x_i. The lines are very similar when wage and wealth is assigned independently or when they are fully correlated. Moreover, the point $\beta=1$, $r=0$ lies on the curve of optimal points as G vanishes in that case. Since we already fixed $r=0.3$ in Subsection 6.4.2, the minimum $\|G(x_{ger})\| = 0.0006$ is attained for $\beta = 1.068$. For the following, we consider the rounded value $\beta = 1.1$ as an appropriate choice, which is also consistent with previous independent estimates of this reinforcement parameter [124, 60].

6.4.4 Simulation results

Figure 6.6 (a) shows the results of simulations with parameters A,n,r,β,γ as specified above and several initial configurations. We simulate with four different initial configurations:

Symmetric:	$X(0) = (1,\dots,1)$
Exponential:	$X_i(0) \sim Exp(1)$ independent
Pareto:	$X_i(0) \sim Pareto(1.5)$ independent with $\mathbb{E}X_i(0) = 1$
γ	$X(0) = A\gamma$



(a) (b)

Figure 6.6: (a) shows $1 - CDF_{sim}$ (6.10) for different initial distributions compared to $1 - CDF_{ger}$ (6.9) (black line) and $1 - CDF$ of scaled wages $228,000\gamma$ (blue line) after $n = 228,000,000$ steps. For symmetric initial condition $X(0) = (1, \dots, 1)$, (b) shows additionally 1-CDF of an intermediate step, scaled to the same mean (light blue). Moreover, the green line shows a stable distribution with $\|G(x)\| = 0$ obtained using Euler's method for the ODE (C.1) starting in $X(n)$ (see Subsection 6.4.5) In all simulations, we took $A = 10,000$, $\beta = 1.1$, $r = 0.3$ and γ as explained in Subsection 6.4.1.

For enhanced comparability, we chose all $X(0)$ such that in each case agents start with one unit of wealth on average. The simulated wealth distribution CDF_{sim} (6.10) after n steps is both compared to the real wealth distribution CDF_{ger} in Germany 2021 (black line) and to the CDF of scaled wages $228,000\gamma$ (blue line), which describes the long-time wealth distribution in a hypothetical world with constant return rates ($\beta = 1$).

Basically, all simulations presented in Figure 6.6 (a) reveal an astonishing accuracy in reproducing CDF_{ger}. The Gini coefficient in our simulations varys between 0.723 (exponential case) and 0.758 (Pareto case) compared the empirical value of 0.75 in Germany 2021 (see [111]). Apart from the net worth of the richest agent, all simulations yield similar wealth distributions, which is consistent with the independence of initial configuration explained in Section 6.2. In fact, the bottom 99% of population is already quite well described by simply re-scaling the wage distribution, since their wealth is mainly determined by savings from wages. Apparently, the impact of increasing returns is negligible for these agents. In contrast to that, the richest percentile reveals a much greater wealth inequality in reality than in the scaled wages distribution, which is well-established for empirical data [34, 112]. The proposed reinforcement mechanism of the Pólya urn model provides an accurate explanation for this significant gap.

Noticeably, the wealth of the richest agent seems to be severely overestimated in the simulation with Pareto distributed initial configuration, whereas it is underestimated in the other simulations. This observation can be underlined by considering the share of total wealth owned by certain parts of the population. The following table shows wealth shares in our simulation in comparison to German data from [111].

Share of richest	50%	10%	1%	0.1%	0.01%
Germany 2021	96.6%	58.9%	28.6%	14.3%	8.17%
Pareto	95.3%	63.1%	34.2%	19.4%	12.4%
Exponential	94.5%	57.9%	24.6%	8.5%	1.4%
Symmetric	94.7%	58.2%	24.9%	8.7%	1.5%
γ	95.6%	61.2%	28.2%	10,8%	2.1%

In fact, this ostensible discrepancy can easily be explained by the enormous variance of wealth within the group of the richest 0.01%. For instance, this group of approximately 7,000 agents starts at a net worth of 40 million Euro and contains 138 billionaires, which own up to 40 billion Euro according to the Forbes List 2021. This billionaire effect has been investigated in detail in [60] for UK data and is already hinted at in the tail behaviour of Figure 1.1. This extremely heterogeneous group is only represented by one agent in our simulation, such that the wealth share of all other agents is significantly impacted by the choice of the representative of the richest group. E.g. in the Pareto case, a rather rich representative of the richest group has been chosen, leading to an overestimation of the wealth share of the richest 0.1%, since this group also contains the representative of the 0.01%, and vice versa for the other simulations. Hence, we should rather consider adjusted wealth shares, where the effect of the richest agent is erased. To be more precise, if $s(\epsilon)$ is the wealth share of the richest $A\epsilon$, $\epsilon \in \{0.1, 0.01, 0.001\}$, agents, then the adjusted share is defined by

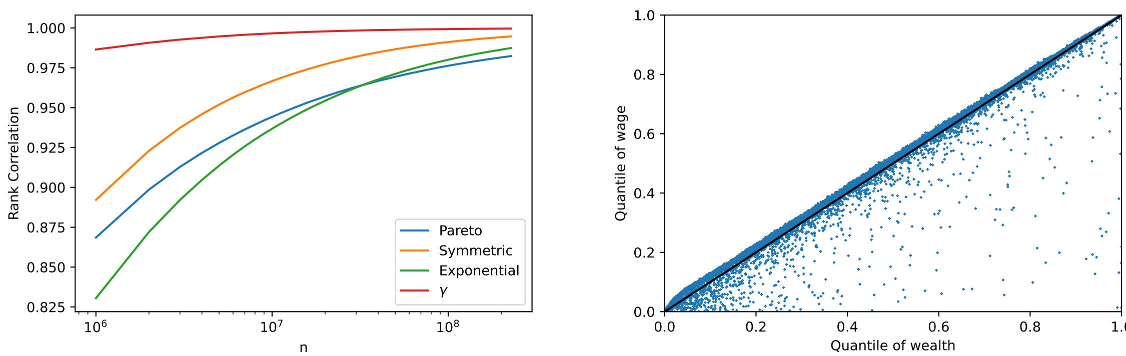
$$s_{ad}(\epsilon) := \frac{s(\epsilon) - s(0.0001)}{1 - s(0.0001)} \,,$$

i.e. s_{ad} is the wealth share, when the richest agent is removed from the system. The following table shows the adjusted wealth shares in our simulations and in reality.

Adjusted share of richest	Germany	Pareto	Exponential	Symmetric	γ
1%	22.2%	24.8 %	23.5%	23.7%	26.6%
0.1%	6.6%	7.9%	7.2%	7.3%	8.8%

With these adjusted shares, we retrieve the good accuracy of all simulations, which we already observed in Figure 6.6.

Correlation between wages and wealth. Figure 6.7 (a) shows the evolution of rank correlation between $X(n)$ and γ. In fact, the rank correlation approaches one in all simulations, even when $X(0)$ is drawn independently of γ. Hence, our process is strongly ordering and the agents with high wage tend to become the richest agents. This feature of our model is opposed to empirical findings, e.g. [34] (or [112]) provide comprehensive statistical information about the common distribution of wealth and income in Germany (or the US) and estimate a much weaker rank correlation of 0.49 (resp. 0.23) between income and net worth. [34] mentions the importance of splitting up wealth through inheritance as a main reason for this moderate correlation, which is not captured by our model. The more-generation model presented in [20] particularly focuses on the impact of inheritance. Moreover, there is an intrinsic difference between [34] and our correlation, because our γ rather represents



(a) (b)

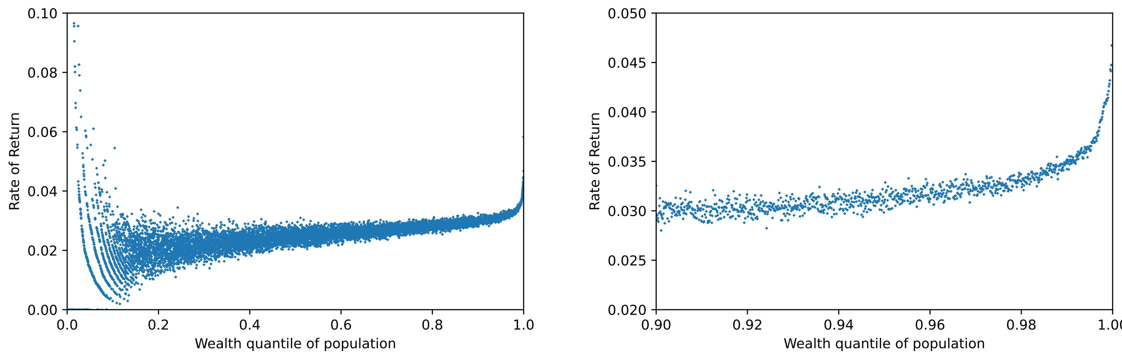
Figure 6.7: The left figure shows the evolution of rank correlation between $X(n)$ and γ for different initial conditions. For the Pareto case, each point in the right figure represents one agent and shows its rank in γ vs. the rank in $X(n), n = 280,000,000$.

savings than income, which obviously increases the correlation with wealth. Furthermore, this overestimation of correlation might emerge as a consequence of the fact that an agent in our model represents a group of 7,000 individuals in reality. Hence, our simulation does not account for untypical individuals that contribute to the moderate empirical rank correlation. Nevertheless, agents, who start with large initial wealth, may keep their advantage in our model for quite a long time as shown in Figure 6.7 (b). Even after $280,000,000$ million steps, there are some agents with only little wage among the richest. This barely occurs for more equal initial configurations. On the other hand, agents cannot have high wages and only little wealth since our process is a pure growth process and wages are assumed to be deterministic.

Increasing returns. As pointed out before, the generalized Pólya urn model with $\beta > 1$ implements the idea of increasing returns, where capital return rates are higher for richer agents. But how does this dependency look like in detail? In order to compute capital return rates in our model, we identify step 216,000,000 with the year 2020, corresponding to an average wealth of 216,327 Euro in 2020. Then we define the rate of return of agent $i \in [A]$ in 2020 as the relative gain in wealth which is not due to wages received,

$$RoR_i := \frac{X_i(228,000,000) - X_i(216,000,000) - 12,000,000 \cdot r\gamma_i}{X_i(216,000,000)} .$$

Figure 6.8 (a) shows these RoR for all agents in the Pareto case, but corresponding plots for the other simulations look essentially the same. For the bottom 10%, we oberserve an enormous variance of the RoR. Whereas many agents had no capital returns at all, some others won one or two steps of the process by luck, leading to large RoRs due to their low level of wealth. This also explains the observable stratified shape of the wealth-RoR-plot for the bottom 20%. For the broad middle class in our simulation, we observe a moderate variance of RoR and only a slight dependence on wealth. Hence, increasing returns mainly concern the



(a) Total population (b) Top 10%

Figure 6.8: Rate of Return on capital between step $n = 216,000,000$ and $n = 228,000,000$ in the Pareto case plotted against wealth quantiles. This corresponds to the year 2020.

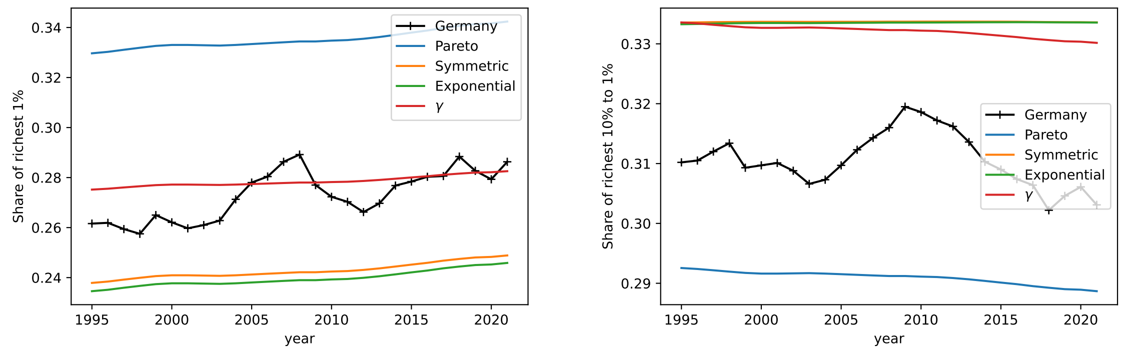
rich. Figure 6.8 (b) shows the RoR of the top 10% in detail. A significant increase of RoR can only be detected for the top 1% of the agents. This is consistent with the observation from Figure 6.6 that the wealth distribution of the bottom 99% is almost equivalent to the "scaled wages"-distribution. Using data from Norway, [53] empirically investigates the dependence of return rates and wealth, which reveals a similar shape of the wealth-return-curve. Moreover, they emphasize that this shape is persistent in time apart from extreme events like the financial crisis, where even decreasing returns could be observed. Almost constant returns for the majority of the population and strongly increasing return rates for the top induce the two-tailed wealth distribution mentioned in Section 6.1.

6.4.5 Time evolution and predictions for the future

Our model seems to be appropriate for describing the present, but does it also reproduce the empirical wealth dynamics of the last decades? And if yes, what does it predict for the future? Since all our simulations are equally valid, we focus in this subsection on the one with symmetric initial condition, which is shown in Figure 6.6 (b). We observe that the wealth distribution after 1/4 of the steps is quite similar to the final distribution, so the attained wealth distribution is fairly stable at least on a moderate time horizon. To identify the steps of our process with years in reality, we consider in this chapter the time changed process

$$t \mapsto Z^{(N)}(t) := \chi\left(\lfloor\left((1+\mu)^t - 1\right) N\rfloor\right) , \tag{6.12}$$

where $t \geq 0$ is the **real time** measured in years and $\mu = \mu(t)$ is the annual growth rate of our economy. For example step $n = 96,000,000$ represents the year 1995 supposing empirical growth rates from [111]. This coincides with an average wealth per agent of approximately 96,000 Euro in 1995. Despite some historical shocks, assuming constant growth $\mu = 0.03$ is



(a) Share of richest 1%

(b) Share of richest 10% to 1%

Figure 6.9: The evolution of the wealth shares of the richest 1% and the following 9% in our simulations compared to German data from [111]. Years have been assigned via (6.12) using empirical μ from [111].

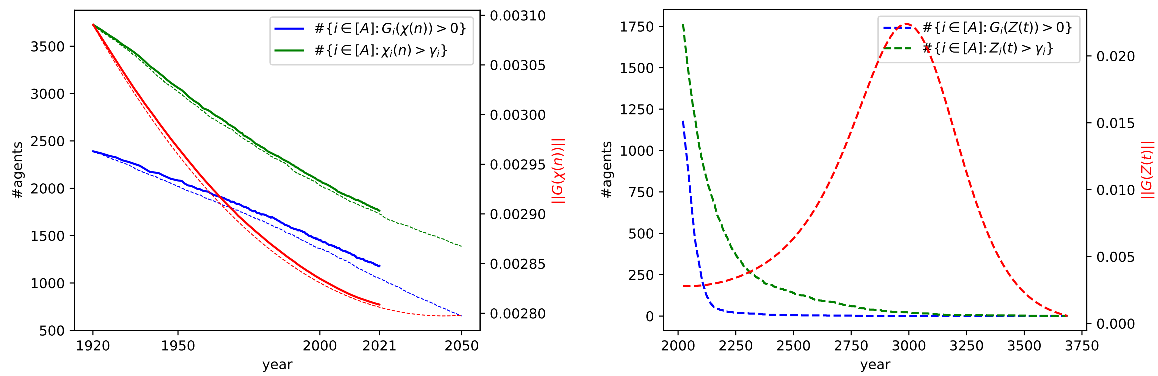
a good approximation over the last 100 years (see Figure C.2), such that we will use this assumption for our future predictions.

Figure 6.9 (a) presents the evolution of the wealth share of the top 1% in our simulations and in reality, all of which reveal a moderate increase. The small differences in the total level of that share have already been discussed above. It should also be noted that the development is much smoother in our simulations than in reality. This is again due to the fact that our model does not encompass the impact of economic shocks. The financial crisis in 2008, for example, led temporarily to a decreasing share of the richest due to falling stock and real estate markets. Figure 6.9 (b) shows the wealth share of the 10-1% quantile, which is almost stagnant with a slightly decreasing trend. Hence, only the richest managed to slowly improve their position over the last decades at the expense of the middle class and even the "moderate" upper class, in reality as well as in our simulations. Consequently, our model did also accurately reproduce the past dynamics of wealth, which justifies to use our model for future predictions.

In Appendix C.1, we present a functional LLN for our process, stating that the process is asymptotically deterministic for large initial values and the dynamics are driven by the field G (6.2), representing the expected increments of wealth shares up to a time scaling (6.3). To be more precise, for large enough N the process $Z^{(N)}$ is well approximated by the solution of the ODE

$$\frac{d}{dt}Z(t) = G(Z(t))\ln(1+\mu) \approx \mu G(Z(t)) \quad \text{with } Z(0) = Z^{(N)}(0)\,. \tag{6.13}$$

This is an efficient tool to make predictions for the future by solving (6.13) with our simulation's result as initial condition, using e.g. Euler's method. Since the number of steps per year in the urn model increases exponentially with (6.12), simulating the model is computationally much more demanding than simply solving (6.13) numerically. Figure 6.10 (a) compares these predictions to our simulations with very good agreement for large initial data. For small or



(a) Simulation

(b) Prediction

Figure 6.10: (a) shows the evolution of number of winners (6.15) and agents with positive $G_i(\chi(n))$ (6.14) as well as $\|G(\chi(n))\|$ in the simulation with $X(0) = (1, \ldots, 1)$. Years are assigned via (6.12) assuming a constant growth rate of $\mu = 0.03$ per year. The dashed lines in (a) show numerical solutions of (6.13) started from simulation data in 1920, which agree well with the time evolution of simulation data. In (b) solutions of (6.13) are shown starting from simulation data in 2021.

moderate initial data fluctuations play a significant role in the stochastic evolution of the urn model. Of course, these predictions suppose that the dynamics of wealth remain unchanged in the future, which might not be the case due to the recent increase of interest rates. We will return to this issue in Section 6.6.

Now, we consider the past and future time evolution of three indicators, presented in Figure 6.10:

The **number of agents with positive field G_i:**

$$\#\{i \in [A] \colon G_i(\chi(n)) > 0\} \tag{6.14}$$

According to (6.13), $G_i(\chi(n))$ is a good indicator for the short-term development of the share of agent $i \in [A]$, where positive (negative) values indicate an increasing (decreasing) share. $G_i(\chi(n)) > 0$ can occur in two different ways: either by large expected capital returns due to large $\chi_i(n)$ or by large wages compared to $\chi_i(n)$. After $n = 280$ million steps (corresponding to the year $t = 2021$), only 11.8% of all agents have positive $G_i(\chi(n)) > 0$. 48% of the richest decentile and even 97% of the richest percentile belong to this group, whereas only 4% of the poorer half of the population do so. Figure C.3 presents a detailed scatter plot of $G_i(\chi(n))$ and $\chi_i(n)$ for two different n. As a consequence, most rich agents will increase their wealth share further, whereas the majority of the population loses. As visible in Figure 6.10 (a), the number of agents with positive short-term trend has been decreasing in time in our simulation. Figure 6.10 (b) presents a long-time prediction for this indicator. The number of agents with positive trend will further decrease until only one agent is left, but this would take another 785 years.

The **number of winners:**

$$\#\{i \in [A] \colon \chi_i(n) > \gamma_i\} \tag{6.15}$$

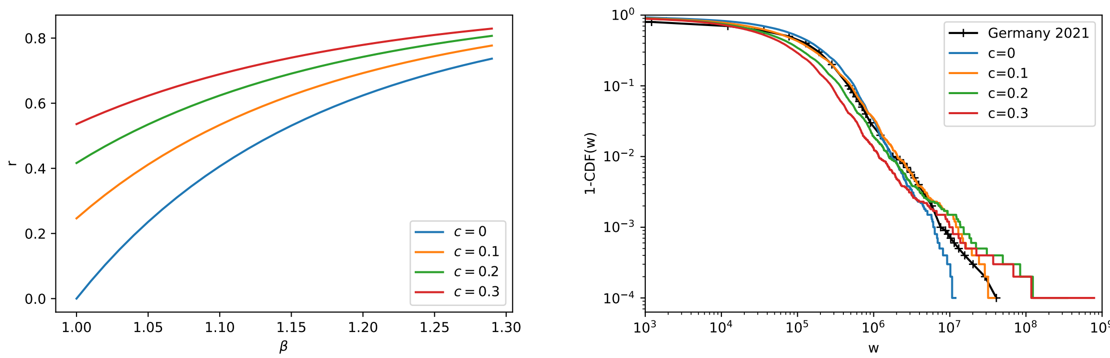
In Section 6.2, we referred to an agent $i \in [A]$ as winner if their wealth share exceeds the wage share, i.e. $\chi_i(n) > \gamma_i$. In that sense, we can identify 18,87% of agents as winners in our simulation in 2021. This group consists exclusively of agents belonging to the bottom 15% or top 5%. The high amount of poor agents in this group is due to the symmetric initial condition, where agents with low wage start with relatively large wealth share. As visible in Figure 6.10, the number of winners has been strongly decreasing in time and will further decrease until only one winner is left, but on an even longer time horizon than the previous indicator.

The Euclidean **norm** $\|G(\chi(n))\|$**.** According to (6.13), $\|G(\chi(n))\|$ can be considered as a measure for the local pace of expected change. In Figure 6.10, we observe decreasing $\|G(\chi(n))\|$ in our simulation and we know from Theorem 6.1 that the norm vanishes asymptotically as we approach a fixed point of G. Following our prediction, it will reach a local minimum in 20 years, followed by a strong increase, which will last over 1,000 years in theory. Finally, it converges exponentially towards zero. Recall that our reinforcement parameter $\beta = 1.1$ was chosen such that $\|G(x_{ger})\|$ is small (c.f. Subsection 6.4.3). In order to gain an intuition for this behaviour, we refer the reader back to the 3-agents case discussed in Figure 6.4 (a). Following a typical trajectory starting near the center of the simplex, it will first approach one of the unstable fixed points, before it finally turns towards a stable fixed point. On this trajectory, the number of winners and agents with $G_i > 0$ is decreasing. Consequently, the observed current local minimum of $\|G(\chi(n))\|$ indicates that our simulated economy is currently close to an unstable fixed point. It may remain near this fixed point for some time, but the dynamics of our model will eventually accelerate and lead the economy into a monopoly-like state.

In this final point, the richest agent dominates the market with a share of 45.8%. Figure 6.6 (b) shows the corresponding **stable wealth distribution**, which is defined analogously to (6.10) for any stable fixed point $x \in \Delta_{A-1}$ (with $228,000x_i$ in the place of $10X_i(n)$). So even within our model, which does not take into account any future changes of parameters or the fundamental mechanism of the dynamics, it would take many centuries until such a monopoly-like state is attained.

6.5 Unequal investment skills as an alternative explanation

Throughout our considerations in Section 6.4, we assumed that return rates on capital do only depend on wealth, but not on individual skills. Nevertheless, it is conceivable that unequal investment skills pose an alternative (or additional) explanation for the gap between wage and wealth distribution (empirically found in e.g. [112, 34]). This question can also be discussed within our extended Pólya urn model. For that, we set $F_i(k) = \alpha_i k^\beta$, where $\alpha_i > 0$ regulates the investment skills of agent i. We keep the parameters $A = 10,000$, $r = 0.3$ and γ as in Section 6.4 since they were derived without using the assumption of equal α_i.



(a) (b)

Figure 6.11: On the left, we see the minimizing $r-\beta$-line as derived in Proposition 6.6 using data from Germany 2021 and sorted γ, but varying c. On the right, we see the stable wealth distributions obtained by Euler's method for (C.1) with $\beta = 1$, $r = 0.3$ and different c. Note that the blue line coincides with the scaled wage distribution.

For specifying α_i, we suppose that there is a positive correlation with wages. For simplicity we use the ansatz

$$\alpha_i = \gamma_i^c \quad \text{for some } c \geq 0\,,$$

where c regulates the intensity of correlation between wages and investment skills. In particular, $c = 0$ corresponds to the the equal skill case from Section 6.4, whereas large c implies huge differences in investment skills. Note that the vector $(\alpha_1, \ldots, \alpha_A)$ does not need to be normalized since only ratios of α_i enter the dynamics (6.1).

In order to find an appropriate β for this situation, we have another look at the $r-\beta$-line derived in Proposition 6.6, which gives the pairs r, β minimizing $\|G(x_{ger})\|$ with a normalized sample x_{ger} from CDF_{ger} (6.9). Hence, we choose again our parameters such that the empirical wealth distribution CDF_{ger} in Germany for 2021 is as close to a stable distribution in our model as possible. Figure 6.11 (a) shows this $r-\beta$-line for several choices of c. First, it is immediately noticeable that positive $r > 0$ is optimal for $\beta = 1$ and any $c > 0$. To get an intuition on this, recall that for $\beta = 1$, $r = 0$ and $c > 0$ our process reveals weak monopoly (see Section 3.2). Since the real wealth distribution is of course more equal than weak monopoly, we need to choose a positive labor share. Second, the optimal labor share is increasing in c. This is due to the fact that larger c increases inequality in our model, which can be compensated by larger r. Third, for fixed r there is an inverse relation between c and β, because they both increase inequality. Note that for large c even $\beta < 1$ is optimal, which corresponds to decreasing return rates. We first focus on a model with $c > 0$ and $\beta = 1$ considering only investment skills and no reinforcement, in order to highlight the conceptual differences to the situation examined in Section 6.4. This case has been rigorously treated in Proposition 6.4, where we proved that the shares $\chi(n)$ converge to a deterministic point and that there are no saddle points for $\beta = 1$ and $c > 0$.

Figure 6.11 (b) shows the unique stable distribution for different c. The model provides a quite good approximation of CDF_{ger} for $c = 0.1$, too. Larger c implies an overestimation of the tail weight, which is consistent with the $r-\beta$-line. Nevertheless, there is a major difference compared to the results presented in Subsection 6.4.4, where the wealth of the richest agent is much larger and more realistic. As visible in Figure 6.6 (right), the model even predicts that the wealth of the richest will further increase in the future for $\beta > 1$, which is not the case for $\beta = 1$, since the distribution in Figure 6.11 (b) is already stable. This can be underlined by a quick look at the wealth shares, as shown in the following table.

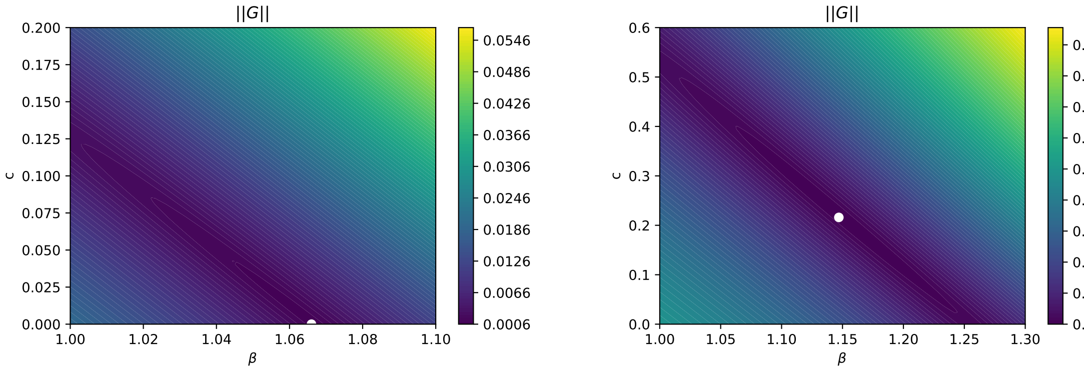
Share of richest	50%	10%	1%	0.1%	0.01%
Germany 2021	96.6%	58.9%	28.6%	14.3%	8.17%
$c = 0.1$	95.4%	59.3%	25.7%	9.4%	1.7%
$c = 0.17$	96.4%	68.0%	40.5%	24.3%	9.0%
$c = 0.2$	96.7%	71.1%	46.5%	31.9%	15.3%
$c = 0.3$	97.3%	77.2%	58.6%	48.0%	22.4%

Indeed, $c = 0.1$ significantly underestimates the wealth of the richest. If we slightly increase c, such that the share of the richest 0.01% coincides with our data, then our model significantly overestimates the 1% and 10% share. Thus, the model with $\beta = 1$ cannot properly reproduce the empirical wealth distribution.

This observation is linked to a conceptual difference between the two models. Whereas the process reveals a random limit in the situation of Subsection 6.4.4 with $\beta > 1$, there is a deterministic limit point for $\beta = 1$, i.e. the long-time limit is fully determined by skills. Hence, under increasing returns with $\beta > 1$ the long time limit is affected by the initial wealth of agents, whereas it is not for constant returns with different investment skills. Moreover, the rank correlation of wage and wealth will reach one after finitely many steps with probability one for $\beta = 1$, $r > 0$, but not necessarily for $\beta > 1$.

It stands to reason that reality is a mixture of both, increasing returns and unequal skills. Figure 6.12 illustrates the goodness of fit for several choices of c and β, again measured by $\|G(x_{ger})\|$ like in Subsection 6.4.3. It underlines the reverse relation between c and β, which means that in the optimum larger β corresponds to smaller c and vice versa. It turns out, that any positive c provides no improvement with respect to this criterion for $r = 0.3$. The reverse relation does also hold for larger r (see Figure 6.12 (b)) since the optimal β-c-line just shifts away from the origin. This is due to the fact that larger c and β increase inequality, but larger r decreases inequality. Nevertheless, for larger r one can achieve slight improvements by taking positive c, since too large β leads to a overestimation of the wealth of the richest agent, in the sense that the losers basically only get their wage. But the total goodness is clearly worse for large r than for our choice $r = 0.3$, supporting that this is a reasonable value.

To summarize the findings in this section, we consider the parameterization of Subsection 6.4.4 as appropriate to model the evolution of the empirical wealth distribution. In particular, the assumption $c = 0$ is well justified and we can ignore differences of investment skills within our model.



(a) $r = 0.3$, $\beta^\star = 1.068$, $c^\star = 0$,
minimal value: $\|G(x_{ger})\| = 0.0006$

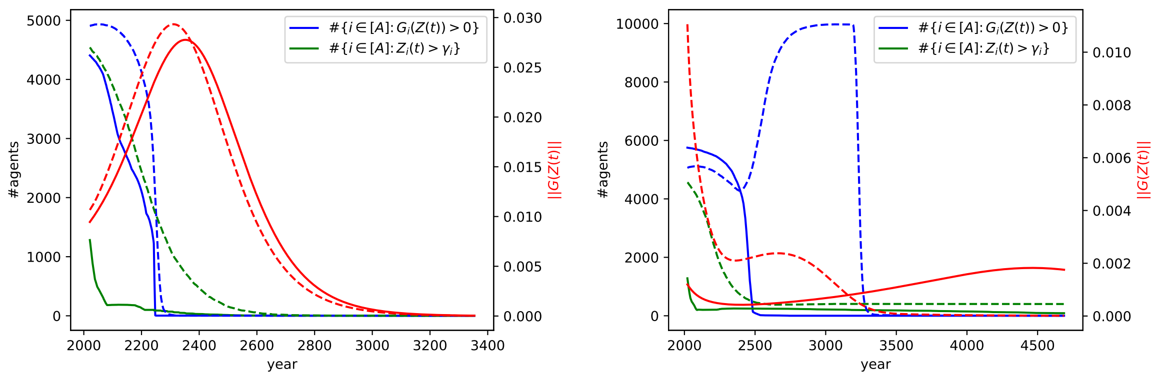
(b) $r = 0.7$, $\beta^\star = 1.147$, $c^\star = 0.216$,
minimal value: $\|G(x_{ger})\| = 0.002$

Figure 6.12: Contour-plot of $\|G(x_{ger})\|$ for varying values of β and c, where $r = 0.3$ resp. $r = 0.7$ and x_{ger} are fixed. The white bullet marks the global minimum ($\beta^\star$, $c^\star$). Note the different scales in both plots.

6.6 Summary and discussion

In Section 6.2, we discussed that our model basically exhibits three different regimes. First, for small $r < r'_c$ there is one random winner, who dominates the population on the long run. All agents have a positive probability of being that winner, which depends on the initial configuration and the wage distribution. Second, for large $r > r_c$, the process converges to a deterministic stable distribution, which is basically a distortion of the wage distribution towards more inequality (for $\beta > 1$). Third, for moderate $r \in (r'_c, r_c)$ there is still a random leading agent, but not all agents can be the leader depending on their wage. To check which regime holds for our choice of parameters we can again compute stable fixed points of the field G by numerically solving (6.13) with different initial conditions. Stability of the generated fixed points was checked with the heuristics from Appendix C.3. Assume w.l.o.g. $\gamma_1 \leq \gamma_2 \leq \ldots$, where $\gamma_1 = 0$ and $\gamma_A = 0.0052$ in our case. Then the solution of (6.13) with initial condition $e^{(1)}$ converges towards a fixed point $(x_1, \ldots, x_A) \in \Delta_{A-1}$ with $x_1 = 0.451$ and $x_A = 0.0042$. Hence, our process with $r = 0.3$ seems to be in the first regime, where even agents with low wage can win the process when they start from a high wealth share. For $r = 0.4$ the numerical solution finds monopoly fixed points for the richest agents but not for the poor, i.e. it converges to a point $x \in \Delta_{A-1}$ with $x_1 = 0$ and $x_A = 0.016$ when it starts in $e^{(1)}$. But with starting point $e^{(A-1)}$, it converges to a point with $x_{A-1} = 0.183$ and $x_A = 0.0074$, corresponding to the monopoly fixed point of agent $A - 1$. Hence, the middle regime applies for $r = 0.4$. Finally, for $r = 0.5$, solutions of (6.13) converge to the same fixed point when starting in $e^{(1)}$ and $e^{(A)}$, such that the process is in the deterministic regime. In this unique stable fixed point $x \in \Delta_{A-1}$ we have $x_1 = 0$ and $x_A = 0.0086$, which is consistent with Proposition 6.2. In summary, we get the rough estimate

$$0.3 < r'_c < 0.4 < r_c < 0.5$$



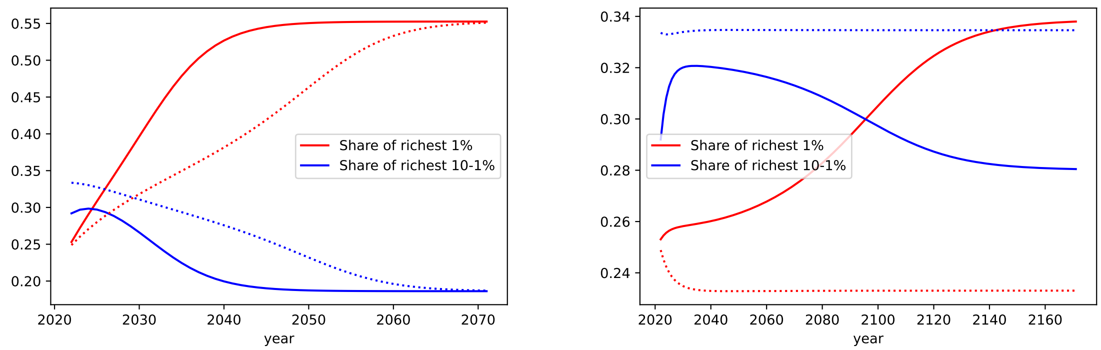
(a) $r = 0.3$

(b) $r = 0.4$

Figure 6.13: Number of winners (green, see (6.15)), agents with positive $G_i(Z(t))$ (blue, see (6.14)) and $\|G(Z(t))\|$ (red) in the solution of (6.13) with initial condition x_{ger} (after normalization) for different labour shares r. For the full (dashed) lines, wage and wealth was assigned fully correlated (uncorrelated). Note the different scales on the time axes. Figure C.4 (a) shows analogously the case $r = 0.5$.

for our situation. As a consequence, even moderately larger $r > 0.3$ can lead to a different regime and long-time behaviour of our process. Recall that our choice $r = 0.3$ is closely linked to a zero-interest economy (see Subsection 6.4.2). Higher interest rates might lead to a larger labor share r and can therefore significantly change our predictions for the future evolution of wealth distribution.

Analogously to Figure 6.10 (b), Figure 6.13 presents the **predictions of our model for different labor share** r, again by solving (6.13). The initial condition is the empirical wealth distribution x_{ger} of 2021, assigned fully correlated with wages or uncorrelated. For the case $r = 0.3 < r'_c$ (Figure 6.13 (a)), we observe basically the same as in Figure 6.10 (b), where we took the final state of our simulation as starting point. The time when the monopoly-like state is attained strongly depends on the wealth of the richest agent, which was underestimated in our simulation (c.f. Subsection 6.4.4). The assignment of wealth and wage is not decisive for the future development as labor plays a minor role. As opposed to that, the predictions do significantly depend on the assignment of wealth and wage for $r = 0.4 \in (r'_c, r_c)$ (Figure 6.13 (b)). For fully correlated assignment, we still observe the monopoly-like behaviour with only one winning agent on the long run, but the dynamics are even slower than for $r = 0.3$. But for initially uncorrelated wealth and wage, we still have several winners with $\chi_i(n) > \gamma_i$ on the long run. Moreover, the number of winners and agents with positive G_i is not monotone in time. This phenomenon does also occur in the $A = 3$ case (see Figure 6.4 (c)), when we follow a trajectory starting near the corner of an agent with low wage. The huge number of agents with positive G_i in the uncorrelated case is due to the effect that the poorest agent starts with a positive share, but converges to zero (Proposition 6.2). Consistent with our theoretical considerations, the limit point is independent of the assignment of wealth and wage



(a) $r = 0.3$ (b) $r = 0.4$

Figure 6.14: Evolution of the wealth share of the richest 1% (red) and the following 9% (blue) in the solution of (6.13) for different labour shares r. We used x_{ger} (after normalization and fully correlated with γ, full line) and the result from the simulation with $X(0) = (1, \ldots, 1)$ (dotted line) as initial condition. Note the different scales on both axes. Figure C.4 (b) shows analogously the case $r = 0.5$

for $r = 0.5 > r_c$, but the way towards this point varies (see Figure C.4 (a)). In the correlated case, the wealth of the richest agent is distributed among all others, which explains the large number of agents with positive G_i. For uncorrelated assignment of wage and wealth, the redistribution is more complex with decreasing number of short-term profiteers. As expected, for $r = 0.4$ and $r = 0.5$ the system converges to a state with more than one winner (green lines). Convergence is slowest in the intermediate regime, where the structure of stable and unstable fixed points is most complex. In all three regimes, the final part of the dynamics is dominated by the fraction of the richest agent, which is the slowest variable in the system.

Using again (6.13), Figure 6.14 presents **predictions for the evolution of wealth shares** in the future. For unchanged labour share $r = 0.3$, we expect that the trend observed in Figure 6.9 will continue for another 30 years, i.e. the richest 1% increase their wealth share to up to 55% and even the moderate upper class loses. Afterwards, the shares remain stable, but there will be a significant redistribution of wealth within the top 1% as indicated in Figure 6.13. In opposition to that, the share of the richest 1-10% stays roughly constant for $r = 0.4$. For $r = 0.4$, we predict only slight changes of the share of the top 1% and the direction of change depends on the chosen starting point. When we even assume $r = 0.5$ for the future, then the richest 1% lose approximately five percentage points of their share to the rest of population within 20 years (see Figure C.4).

However, as discussed before, the actual stable points of the dynamics may not be reached in reality for various reasons. While the structure of stable limit distributions changes over time due to external influences that are not included in our model, the system evolves only slowly in a complex landscape with many fixed points with unstable directions. We have seen in Figures 6.10 and Figure 6.13 that it would take hundreds of years to reach the stable distribution with model parameters fitted from today's data.

In summary, the proposed model provides an accurate replication of the observed wealth distribution in Germany given the distribution of wages, widely independently of the presumed initial wealth distribution. In particular, the two tailed structure from Figure 1.1 is well reproduced. There is only some discrepancy concerning the wealth of the richest agent in our model, who represents the richest 0.01% of real population. Since there is a huge variance within the wealth of this group, we would have to simulate single households to properly represent this group. This would be computationally much more costly for a rather limited gain, and we consider this only a minor disadvantage. Moreover, the observed wealth dynamics of recent decades is reflected in our simulation, where the wealth share of the richest percent of population grows slightly at the expense of the rest, even the "moderate" upper class. According to our model, this trend will continue in the future and less and less people will profit from increasing returns. The return rates on capital are also accurately modeled by the Pólya urn mechanism, implying that increasing returns do basically only affect the richest percentile (which eventually leads to the two-tailed structure of wealth distribution).

In order to understand the generic nature of wealth dynamics and avoid overfitting, the model is kept intentionally very simple, which naturally leaves space for further refinements and research questions: A major intrinsic disadvantage of the proposed model is that it is strongly ordering, i.e. the rank correlation between wage and capital is close to one on the long run. This does not comply with empirical data. Including inhomogeneous investment skills (represented by the parameter α) might pose a solution to this problem, if investment skills and wages are not chosen fully correlated. Due to a lack of useful data on the correlation structure of wealth and wage, we leave this issue open for future research. We concluded that unequal investment skills with constant return rates provide a less accurate explanation of empirical observations, but more refined research could be done here, including a more sophisticated model for the fitness of agents.

Chapter 7

Conclusions and Open Problems

When my supervisor first suggested this topic in February 2022, it was originally intended as a "short" starting project. The goal was to find a fairly general criterion to predict the strong monopolist in the limit for large initial market size, which allows a discussion of the dependence of long-time limits on the fitness of agents. Soon, the task proved to be slightly more involved than expected. As a consequence, we had to distinguish different classes of feedback functions, which we later called type P and E. Diving into the vast literature on generalized Pólya urns, we realized that many interesting results (like Theorem 2.6 on the wealth of losers) are subject to restrictive assumptions, which did not seem to be necessary. Moreover, we did not find a suitable scaling limit for the dynamics of our specific model or a simple formula to compute the deterministic limit point in the sub-linear case. Furthermore, a detailed description of the transition between the deterministic and the monopoly regime seemed not to have been studied before. Hence, we were left with many open problems to resolve in the following two years, and the answers we found resulted in this work. What I found particularly astonishing (and new in this context) is the emergence of random weak monopoly (Section 3.3) for almost linear feedback and the "loser paradox" in Chapter 5. After studying the theory of non-linear Pólya urns, we were excited to finally examine in how far this model helps to understand phenomena from the real world, like the emergence of inequality, which led to further fascinating insights (Chapter 6).

In spite of having added more than a hundred pages to the literature of Pólya urns, we can still find space for some (maybe minor) refinements. These are for example:

1. In our main result on strong monopoly (Theorem 3.5), we excluded feedback close to the identity via assumption (3.4). Nevertheless, Example 3.19 indicates that it is still possible to identify attraction domains when (3.4) is not satisfied. Hence, an extension of Theorem 3.5 seems possible. Moreover, in the sub-linear case, Corollary 3.26 excludes exponentially decaying feedback, but Appendix A hints that Condition (C2) can be weakened (but not omitted).

2. In the case of feedback close to the identity, Corollary 3.40 derives a lower bound for the rate of convergence $\chi(n) \to \chi(\infty)$, which turns out to be sharp in the situation of Example 3.41. One could try to find a general asymptotic here. For general feedback

in Section 4.1, we found that time of order N is not sufficient for the process to escape from an unstable equilibrium. Quantifying this time, however, is pending.

3. In Subsection 5.2.2 and Appendix D.3, we mentioned that the constant $c(A, a)$, which is a measure for the tail dependence of losing agents, seems to be increasing with the number of agents. Nevertheless, this surprising observation remains to be proven (Conjecture D.9).

4. In Chapter 6, we heuristically discussed the structure of the zero-set of an extended model, which could be completed by a formal argument, in particular for moderate labour share r. Moreover, refining the investigations from Section 6.5 on unequal investment skills as an alternative explanation for the gap between wage and wealth distribution could be enriching.

5. Appendix E derives a result on the time of monopoly as an application of Theorem 5.10 and Theorem 5.16. A small gap in the proof (see (E.1)), though, still has to be filled. In addition, an extension to general feedback would be desirable.

Many more unsolved questions arise in the recent field of Pólya urns with infinitely many colours, which are mostly implemented as measure-valued urn models [10, 93, 9, 80]. So far, most literature has focused on the extension of linear replacement mechanisms to the infinite colour case, but it might be interesting to consider the non-linear case, too.

In summary, even 100 years after Pólya and Eggenberger [52] initially proposed this model, it still is an almost inexhaustible source of new questions, extensions and applications.

Every one to whom much is given, of him will much be required; and of him to whom men commit much they will demand the more.

- Gospel of Luke, 12:48, RSV

Bibliography

[1] Mohamed S. Abdel-Hameed and Frank Proschan. "Shock models with underlying birth process". In: *Journal of Applied Probability* 12.1 (1975), pp. 18–28 (cit. on p. 69).

[2] Giacomo Aletti and Irene Crimaldi. "Generalized Rescaled Pólya urn and its statistical application". In: *Electronic Journal of Statistics* 16.1 (2022), pp. 1635–1680 (cit. on p. 2).

[3] Giacomo Aletti and Irene Crimaldi. "The Rescaled Pólya Urn and the Wright—Fisher Process with Mutation". In: *Mathematics* 9.22 (2021), p. 2909 (cit. on p. 2).

[4] Giacomo Aletti and Irene Crimaldi. "The rescaled Pólya urn: local reinforcement and chi-squared goodness-of-fit test". In: *Advances in Applied Probability* 54.3 (2022), pp. 849–879 (cit. on p. 2).

[5] W. Brian Arthur et al. *Increasing returns and path dependence in the economy.* University of michigan Press, 1994 (cit. on pp. 1–3, 96).

[6] W. Brian Arthur, Yu M. Ermoliev, and Yu M. Kaniovski. "Limit theorems for proportions of balls in a generalized urn scheme". In: *Working paper IIASA WP-87-111* (1987) (cit. on pp. 12, 55).

[7] Laurent Bach, Laurent E. Calvet, and Paolo Sodini. "Rich pickings? Risk, return, and skill in the portfolios of the wealthy". In: *HEC Research Papers Series 1126* (2015) (cit. on p. 96).

[8] Zhi-Dong Bai, Feifang Hu, and Li-Xin Zhang. "Gaussian approximation theorems for urn models and their applications". In: *The Annals of Applied Probability* 12.4 (2002), pp. 1149–1173 (cit. on p. 55).

[9] Antar Bandyopadhyay and Debleena Thacker. "A new approach to Pólya urn schemes and its infinite color generalization". In: *The Annals of Applied Probability* 32.1 (2022), pp. 46–79 (cit. on pp. 2, 126).

[10] Antar Bandyopadhyay and Debleena Thacker. "Pólya urn schemes with infinitely many colors". In: *Bernoulli* 23.4B (2017), pp. 3243–3267 (cit. on pp. 2, 126).

[11] Andrew D. Barbour. "The asymptotic behaviour of birth and death and some related processes". In: *Advances in Applied Probability* 7.1 (1975), pp. 28–43 (cit. on pp. 40, 70).

[12] Néstor R. Barraza. "A new homogeneous pure birth process based software reliability model". In: *Proceedings of the 38th International Conference on Software Engineering Companion*. 2016, pp. 710–712 (cit. on p. 69).

[13] Maurice S. Bartlett. *An introduction to stochastic processes*. University Press Cambridge, 1966 (cit. on p. 70).

[14] Alessandro Bellina, Giordano De Marzo, and Vittorio Loreto. "Time-Dependent Urn Models reproduce the full spectrum of novelties discovery". In: *arXiv preprint arXiv:2401.10114* (2024) (cit. on p. 2).

[15] Michel Benaim. "A dynamical system approach to stochastic approximations". In: *SIAM Journal on Control and Optimization* 34.2 (1996), pp. 437–472 (cit. on p. 50).

[16] Michel Benaim. *Seminaire de probabilites XXXIII*. Springer, 2006, pp. 1–68 (cit. on p. 50).

[17] Michel Benaim, Sebastian J. Schreiber, and Pierre Tarres. "Generalized urn models of evolutionary processes". In: *Annals of Applied Probability* 14.3 (2004), pp. 1455–1478 (cit. on p. 50).

[18] Michel Benaim et al. "A generalized Pólya's urn with graph based interactions". In: *Random Structures & Algorithms* 46.4 (2015), pp. 614–634 (cit. on pp. 3, 98).

[19] Jess Benhabib and Alberto Bisin. "Skewed wealth distributions: Theory and empirics". In: *Journal of Economic Literature* 56.4 (2018), pp. 1261–91 (cit. on pp. 1, 95, 96).

[20] Jess Benhabib, Alberto Bisin, and Shenghao Zhu. "The distribution of wealth and fiscal policy in economies with finitely lived agents". In: *Econometrica* 79.1 (2011), pp. 123–157 (cit. on pp. 96, 113).

[21] Jean Bertoin. "Limits of Pólya urns with innovations". In: *Electronic Journal of Probability* 28 (2023), pp. 1–19 (cit. on p. 2).

[22] Bundeszentrale für politische Bildung. *Sparverhalten nach Einkommen*. Accessed on July 18, 2023. URL: `https://www.bpb.de/system/files/dokument_pdf/08%20Sparverhalten.pdf` (cit. on p. 109).

[23] David Blackwell and David Kendall. "The Martin boundary for Polya's urn scheme, and an application to stochastic population growth". In: *Journal of Applied Probability* 1.2 (1964), pp. 284–296 (cit. on p. 9).

[24] David Blackwell and James B. MacQueen. "Ferguson Distributions Via Polya Urn Schemes". In: *The Annals of Statistics* 1.2 (1973), pp. 353–355 (cit. on pp. 2, 9).

[25] Bruce M. Boghosian. "Kinetics of wealth and the Pareto law". In: *Physical Review E* 89.4 (2014) (cit. on p. 96).

[26] Dennis D. Boos. "A converse to Scheffe's theorem". In: *The Annals of Statistics* 13.1 (1985), pp. 423–427 (cit. on p. 73).

[27] Vivek S. Borkar. *Stochastic approximation: a dynamical systems viewpoint*. Vol. 48. Springer, 2009 (cit. on pp. 13, 49, 50, 55).

[28] Konstantin Borovkov. "Gaussian process approximations for multicolor Pólya urn models". In: *Journal of Applied Probability* 58.1 (2021), pp. 274–286 (cit. on pp. 12, 55).

[29] Jean-Philippe Bouchaud and Marc Mézard. "Wealth condensation in a simple model of economy". In: *Physica A: Statistical Mechanics and its Applications* 282.3-4 (2000), pp. 536–545 (cit. on p. 96).

[30] W. Brian Arthur, Yu M. Ermoliev, and Yu M. Kaniovski. "Strong laws for a class of path-dependent stochastic processes with applications". In: *Stochastic optimization*. Springer, 1986, pp. 287–300 (cit. on pp. 3, 10, 33, 98, 99).

[31] Statistisches Bundesamt. *Lohn- und Einkommenssteuer 2018*. Accessed on April 18, 2023. URL: https://www.destatis.de/DE/Themen/Staat/Steuern/Lohnsteuer-Einkommensteuer/Publikationen/Downloads-Lohn-und-Einkommenssteuern/lohn-einkommensteuer-2140710187004.pdf?__blob=publicationFile (cit. on pp. 108, 146).

[32] Statistisches Bundesamt. *Volkswirtschaftliche Gesamtrechnungen 2022*. Accessed on Juni 21, 2023. URL: https://www.destatis.de/DE/Themen/Wirtschaft/Volkswirtschaftliche-Gesamtrechnungen-Inlandsprodukt/Publikationen/Downloads-Inlandsprodukt/inlandsprodukt-lange-reihen-pdf-2180150.pdf?__blob=publicationFile (cit. on pp. 109, 146).

[33] Bundesfinanzministerium. *BMF Monatsbericht Januar 2022*. Accessed on Juni 28, 2023. URL: https://www.bundesfinanzministerium.de/Monatsberichte/2022/01/Inhalte/Kapitel-6-Statistiken/6-4-04-einkommensverteilung.html (cit. on pp. 109, 146).

[34] Bundesministerium für Arbeit und Soziales. *Analyse der Verteilung von Einkommen und Vermögen in Deutschland*. Accessed on Juni 24, 2023. URL: https://www.armuts-und-reichtumsbericht.de/SharedDocs/Downloads/Service/Studien/analyse-verteilung-einkommen-vermoegen.pdf?__blob=publicationFile&v=3 (cit. on pp. 95, 112, 113, 118).

[35] Ben-Hur Francisco Cardoso, Sebastián Gonçalves, and José Roberto Iglesias. "Wealth distribution models with regulations: Dynamics and equilibria". In: *Physica A: Statistical Mechanics and its Applications* 551 (2020) (cit. on p. 96).

[36] James A. Cavender. "Quasi-Stationary Distributions of Birth-and-Death Processes". In: *Advances in Applied Probability* 10.3 (1978), pp. 570–586 (cit. on p. 78).

[37] Bikas K. Chakrabarti et al. *Econophysics of income and wealth distributions*. Cambridge University Press, 2013 (cit. on p. 95).

[38] Arnab Chatterjee, Bikas K. Chakrabarti, and Subhrangshu S. Manna. "Pareto law in a kinetic model of market with random saving propensity". In: *Physica A: Statistical Mechanics and its Applications* 335.1-2 (2004), pp. 155–163 (cit. on pp. 95, 96, 109).

[39] Dimitris Cheliotis and Dimitra Kouloumpou. "Functional limit theorems for the Polya and q-Polya urns". In: *arXiv preprint arXiv:1905.13336* (2019) (cit. on pp. 12, 55).

[40] Fan Chung, Shirin Handjani, and Doug Jungreis. “Generalizations of Polya’s urn problem”. In: *Annals of combinatorics* 7.2 (2003), pp. 141–153 (cit. on p. 3).

[41] Andrea Collevecchio, Codina Cotar, and Marco LiCalzi. “On a preferential attachment and generalized Pólya’s urn model”. In: *The Annals of Applied Probability* 23.3 (2013), pp. 1219–1253 (cit. on p. 2).

[42] Don Coppersmith and Persi Diaconis. *Random walk with reinforcement.* Unpublished manuscript, 1986 (cit. on p. 2).

[43] Marcelo Costa and Jonathan Jordan. “Phase transitions in non-linear urns with interacting types”. In: *Bernoulli* 28.4 (2022), pp. 2546–2562 (cit. on pp. 3, 4, 54).

[44] Codina Cotar and Vlada Limic. “Attraction time for strongly reinforced walks”. In: *The Annals of Applied Probability* 19.5 (2009) (cit. on pp. 2, 11, 13, 69, 81, 88, 163, 168, 169).

[45] Burgess Davis. “Reinforced random walk”. In: *Probability Theory and Related Fields* 84.2 (1990), pp. 203–229 (cit. on pp. 2, 7, 8).

[46] Christopher B.C. Dean. “Functional limit theorems for Pólya urns with growing initial compositions”. In: *arXiv preprint arXiv:2206.05138* (2022) (cit. on p. 55).

[47] Elosa Diaz-Frances and Luis G. Gorostiza. “Inference and model comparison for species accumulation functions using approximating pure birth processes”. In: *Journal of Agricultural, Biological, and Environmental Statistics* 7 (2002), pp. 335–349 (cit. on p. 69).

[48] Adrian Drăgulescu and Victor M. Yakovenko. “Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States”. In: *Physica A: Statistical Mechanics and its Applications* 299.1-2 (2001), pp. 213–221 (cit. on p. 95).

[49] Eleni Drinea, Alan Frieze, and Michael Mitzenmacher. “Balls and bins models with feedback”. In: *SODA*. Vol. 2. Citeseer. 2002, pp. 308–315 (cit. on pp. 3, 32).

[50] Joe Dunlop. “Monopoly in Balls-in-Bins Processes with Reinforcement and Fitness”. MA thesis. University of Warwick, 2020 (cit. on pp. 11, 29, 32).

[51] Stefan Ederer, Maximilian Mayerhofer, and Miriam Rehm. “Rich and ever richer? Differential returns across socioeconomic groups”. In: *Journal of Post Keynesian Economics* 44.2 (2021), pp. 283–301 (cit. on p. 96).

[52] Florian Eggenberger and George Pólya. “Über die Statistik verketteter Vorgänge”. In: *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 3.4 (1923), pp. 279–289 (cit. on pp. 2, 3, 126).

[53] Andreas Fagereng et al. “Heterogeneity and persistence in returns to wealth”. In: *Econometrica* 88.1 (2020), pp. 115–170 (cit. on pp. 1, 96, 115).

[54] William Feller. *An introduction to probability theory and its applications.* Wiley & Sons, 1947 (cit. on p. 8).

[55] William Feller. “Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung”. In: *Acta Biotheoretica* 5.1 (1939), pp. 11–40 (cit. on p. 70).

[56] William Feller. "On the integro-differential equations of purely discontinuous Markoff processes". In: *Transactions of the American Mathematical Society* 48.3 (1940), pp. 488–515 (cit. on p. 70).

[57] Shui Feng. *The Poisson-Dirichlet distribution and related topics: models and asymptotic behaviors*. Springer Science & Business Media, 2010 (cit. on pp. 156, 157).

[58] Stephen Fletcher. "Nucleation on active sites: Part III. Nucleation modelled as a pure birth process and nucleation modelled as a birth-and-death process". In: *Journal of electroanalytical chemistry and interfacial electrochemistry* 215.1-2 (1986), pp. 1–9 (cit. on p. 69).

[59] Samuel Forbes. "A Study of the Probability Distribution of the Balls in Bins Process with Power Law Feedback". In: *arXiv preprint arXiv:2308.10734* (2023) (cit. on p. 70).

[60] Samuel Forbes and Stefan Grosskinsky. "A study of UK household wealth through empirical analysis and a non-linear Kesten process". In: *Plos one* 17.8 (2022) (cit. on pp. 3, 95, 96, 111, 113).

[61] Simone Franchini. "Large deviations for generalized Pólya urns with arbitrary urn function". In: *Stochastic Processes and their Applications* 127.10 (2017), pp. 3372–3411 (cit. on p. 12).

[62] David A. Freedman. "Bernard Friedman's Urn". In: *The Annals of Mathematical Statistics* 36.3 (1965), pp. 956–970 (cit. on p. 9).

[63] Bernard Friedman. "A simple urn model". In: *Communications on Pure and Applied Mathematics* 2.1 (1949), pp. 59–70 (cit. on p. 2).

[64] Wendell H. Furry. "On fluctuation phenomena in the passage of high energy electrons through lead". In: *Physical Review* 52.6 (1937), p. 569 (cit. on p. 69).

[65] Xavier Gabaix. "Power laws in economics and finance". In: *Annu. Rev. Econ.* 1.1 (2009), pp. 255–294 (cit. on pp. 95, 96).

[66] Larry Goldstein and Gesine Reinert. "Stein's method for the Beta distribution and the Pólya-Eggenberger urn". In: *Journal of Applied Probability* 50.4 (2013), pp. 1187–1205 (cit. on p. 9).

[67] Thomas Gottfried and Stefan Grosskinsky. "Asymptotics of generalized Pólya urns with non-linear feedback". In: *Electron. J. Probab.* 29.92 (2024), pp. 1–56 (cit. on pp. 4, 15, 49).

[68] Thomas Gottfried and Stefan Grosskinsky. "Tails of explosive birth processes and applications to non-linear Pólya urns". In: *arXiv preprint arXiv:2406.15006* (2024) (cit. on pp. 4, 69).

[69] Thomas Gottfried and Stefan Grosskinsky. "Wages and Capital returns in a generalized Pólya urn". In: *arXiv preprint arXiv:2401.17688* (2024) (cit. on pp. 4, 95).

[70] Raul Gouet. "Martingale functional central limit theorems for a generalized Pólya urn". In: *The Annals of Probability* (1993), pp. 1624–1639 (cit. on p. 55).

[71] Xiaoying Han and Peter E. Kloeden. *Random ordinary differential equations and their numerical solution.* Springer, 2017 (cit. on p. 64).

[72] Norbert Henze. *Stochastik: eine Einführung mit Grundzügen der Maßtheorie: inkl. zahlreicher Erklärvideos.* Springer-Verlag, 2019 (cit. on p. 35).

[73] Bruce M. Hill, David Lane, and William Sudderth. "A strong law for some generalized urn processes". In: *The Annals of Probability* 8.2 (1980), pp. 214–226 (cit. on pp. 2, 3, 9).

[74] Bruce M. Hill, David Lane, and William Sudderth. "Exchangeable urn processes". In: *The Annals of Probability* (1987), pp. 1586–1592 (cit. on p. 9).

[75] Fred M. Hoppe. "Pólya-like urns and the Ewens' sampling formula". In: *Journal of Mathematical Biology* 20.1 (1984), pp. 91–94 (cit. on p. 2).

[76] Zhishui Hu and Yiting Zhang. "Strong limit theorems for step-reinforced random walks". In: *arXiv preprint arXiv:2311.15263* (2023) (cit. on p. 96).

[77] Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes.* Vol. 288. Springer Science & Business Media, 2013 (cit. on pp. 50, 65).

[78] Ian R. James and James E. Mosimann. "A New Characterization of the Dirichlet Distribution Through Neutrality". In: *The Annals of Statistics* 8.1 (1980), pp. 183–189 (cit. on pp. 9, 42).

[79] Svante Janson. "Functional limit theorems for multitype branching processes and generalized Pólya urns". In: *Stochastic Processes and their Applications* 110.2 (2004), pp. 177–245 (cit. on p. 55).

[80] Svante Janson, Cecile Mailler, and Denis Villemonais. "Fluctuations of balanced urns with infinitely many colours". In: *Electronic Journal of Probability* 28 (2023), pp. 1–72 (cit. on pp. 2, 126).

[81] Bo Jiang et al. "On the duration and intensity of competitions in nonlinear Pólya urn processes with fitness". In: *ACM SIGMETRICS Performance Evaluation Review* 44.1 (2016), pp. 299–310 (cit. on pp. 3, 12, 32).

[82] Norman L. Johnson and Samuel Kotz. *Urn models and their application: an approach to modern discrete probability theory.* Wiley, 1977 (cit. on p. 3).

[83] Michael J. Kearney and Richard J. Martin. "First passage properties of a generalized Pólya urn". In: *Journal of Statistical Mechanics: Theory and Experiment* 2016.12 (2016), p. 123407 (cit. on p. 3).

[84] Kostya Khanin and Raya Khanin. "A probabilistic model for the establishment of neuron polarity". In: *Journal of Mathematical Biology* 42.1 (2001), pp. 26–40 (cit. on pp. 3, 12, 32).

[85] Gustavo Kohlrausch and Sebastián Gonçalves. "Wealth distribution on a dynamic complex network". In: *arXiv preprint arXiv:2302.03677* (2023) (cit. on p. 96).

[86] Konrad Kolesko and Ecaterina Sava-Huss. "Gaussian fluctuations for the two urn model". In: *arXiv preprint arXiv:2301.08602* (2023) (cit. on p. 3).

[87] Dimitrios G. Konstantinides and Charalampos D. Passalidis. “A new approach in two-dimensional heavy-tailed distributions”. In: *arXiv preprint arXiv:2402.09040* (2024) (cit. on p. 83).

[88] Samuel Kotz and Narayanaswamy Balakrishnan. “Advances in Urn Models During the Past Two”. In: *Advances in Combinatorial Methods and Applications to Probability and Statistics* (2012), p. 203 (cit. on p. 3).

[89] Arnaud de La Fortelle. “Yule process sample path asymptotics”. In: *Elect. Comm. in Probab.* 11 (2006), pp. 193–199 (cit. on p. 70).

[90] Kang K.L. Liu et al. “Simulation of a generalized asset exchange model with economic growth and wealth distribution”. In: *Physical Review E* 104.1 (2021), p. 014150 (cit. on p. 96).

[91] Ove Lundberg. “On random processes and their application to sickness and accident statistics”. PhD thesis. Stockholm, Sweden, 1940 (cit. on p. 70).

[92] Hosam Mahmoud. *Pólya urn models.* CRC press, 2008 (cit. on p. 3).

[93] Cécile Mailler and Jean-François Marckert. “Measure-valued Pólya urn processes”. In: *Electronic Journal of Probability* 22 (2017), pp. 1–33 (cit. on pp. 2, 126).

[94] Andreı A. Markov. “Sur quelques formules limites du calcul des probabilités”. In: *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* 11.3 (1917), pp. 177–186 (cit. on p. 2).

[95] Servet Martínez. “Quasi-Stationary Distributions for Birth-Death Chains. Convergence Radii and Yaglom Limit”. In: *Cellular Automata and Cooperative Systems.* Springer Netherlands, 1993, pp. 491–505 (cit. on p. 78).

[96] Krishanu Maulik and Manit Paul. “Feedback Interacting Urn Models”. In: *arXiv preprint arXiv:2211.07573* (2022) (cit. on p. 3).

[97] Mikhail Menshikov and Vadim Shcherbakov. “Balls-in-bins models with asymmetric feedback and reflection”. In: *arXiv preprint arXiv:2204.05724* (2022) (cit. on pp. 3, 11, 12, 32).

[98] Michael Mitzenmacher, Roberto Oliveira, and Joel Spencer. “A scaling result for explosive processes”. In: *the electronic journal of combinatorics* 11.1 (2004), R31 (cit. on pp. 3, 10, 32).

[99] Andrea Monaco, Matteo Ghio, and Adamaria Perrotta. “Wealth dynamics in a multi-aggregate closed monetary system”. In: *Physica A: Statistical Mechanics and its Applications* (2024), p. 129851 (cit. on p. 96).

[100] Mikhail B. Nevel’son and Rafail Z. Has’ minskii. *Stochastic approximation and recursive estimation.* Vol. 47. American Mathematical Soc., 1976 (cit. on pp. 13, 49, 98, 99).

[101] Roberto Oliveira. “Balls-in-bins processes with feedback and brownian motion”. In: *Combinatorics, Probability and Computing* 17.1 (2008), pp. 87–110 (cit. on pp. 3, 12, 16).

[102] Roberto Oliveira. "The onset of dominance in balls-in-bins processes with feedback". In: *Random Structures & Algorithms* 34.4 (2009), pp. 454–477 (cit. on pp. 3, 10, 11, 16, 69, 81).

[103] Roberto Oliveira and Joel Spencer. "Avoiding defeat in a balls-in-bins process with feedback". In: *arXiv preprint math/0510663* (2005) (cit. on pp. 3, 12, 32).

[104] Enzo Orsingher and Federico Polito. "Randomly stopped nonlinear fractional birth processes". In: *Stochastic Analysis and Applications* 31.2 (2013), pp. 262–292 (cit. on p. 70).

[105] Anthony G. Pakes. "Divergence rates for explosive birth processes". In: *Stochastic processes and their applications* 41.1 (1992), pp. 91–99 (cit. on p. 70).

[106] Vilfredo Pareto. *Cours d'économie politique*. F. Rouge, 1896 (cit. on p. 95).

[107] Robin Pemantle. "A survey of random processes with reinforcement". In: *Probability surveys* 4 (2007), pp. 1–79 (cit. on pp. 2, 13, 49, 97, 98).

[108] Robin Pemantle. "Nonconvergence to unstable points in urn models and stochastic approximations". In: *The Annals of Probability* 18.2 (1990), pp. 698–712 (cit. on p. 99).

[109] Thomas Piketty. *A Brief History of Equality*. Harvard University Press, 2022 (cit. on p. 1).

[110] Thomas Piketty. *Capital in the 21st Century*. Harvard University Press, 2017 (cit. on p. 1).

[111] Thomas Pikkety. *World Inequality Database*. Accessed on April 18, 2023. URL: `https://wid.world/data/` (cit. on pp. 2, 107, 109, 112, 115, 116, 146, 147).

[112] Vincenzo Quadrini and José-Vıctor Rıos-Rull. "Dimensions of inequality: Facts on the US distribution of earnings, income and wealth". In: *Federal Reserve Bank of Minneapolis Quarterly Review* 21.2 (1997), pp. 3–21 (cit. on pp. 95, 112, 113, 118).

[113] Herbert Robbins and Sutton Monro. "A stochastic approximation method". In: *The annals of mathematical statistics* (1951), pp. 400–407 (cit. on p. 49).

[114] Ioanid Roşu and Fahad Saleh. "Evolution of shares in a proof-of-stake cryptocurrency". In: *Management Science* 67.2 (2021), pp. 661–672 (cit. on p. 2).

[115] Wioletta M. Ruszel and Debleena Thacker. "Positive reinforced generalized time-dependent Pólya urns via stochastic approximation". In: *arXiv preprint arXiv:2201.12603* (2022) (cit. on p. 3).

[116] Paul A. Samuelson. "A fallacy in the interpretation of Pareto's law of alleged constancy of income distribution". In: *Rivista Internazionale di Scienze Economiche e Commerciali* 12 (1965), pp. 246–250 (cit. on p. 1).

[117] Hristo Sariev, Sandra Fortini, and Sonia Petrone. "Infinite-color randomly reinforced urns with dominant colors". In: *Bernoulli* 29.1 (2023), pp. 132–152 (cit. on p. 2).

[118] Sebastian J. Schreiber. "Urn models, replicator processes, and random genetic drift". In: *SIAM Journal on Applied Mathematics* 61.6 (2001), pp. 2148–2167 (cit. on p. 50).

[119] Qin Shuo. “Interacting urn models with strong reinforcement”. In: *arXiv preprint arXiv2311.13480* (2023) (cit. on p. 3).

[120] Herbert A. Simon. “On a class of skew distribution functions”. In: *Biometrika* 42.3/4 (1955), pp. 425–440 (cit. on p. 96).

[121] Somya Singh, Fady Alajaji, and Bahman Gharesifard. “Generating preferential attachment graphs via a Pólya urn with expanding colors”. In: *Network Science* 12.2 (2024), pp. 139–159 (cit. on p. 3).

[122] Wenpin Tang. “Stability of shares in the Proof of Stake Protocol–Concentration and Phase Transitions”. In: *arXiv preprint arXiv:2206.02227* (2022) (cit. on p. 2).

[123] Bálint Tóth. “Limit theorems for weakly reinforced random walks on Z”. In: *Studia Scientiarum Mathematicarum Hungarica* 33.1 (1997), pp. 321–338 (cit. on p. 2).

[124] Hunter A. Vallejos, James J. Nutaro, and Kalyan S. Perumalla. “An agent-based model of the observed distribution of wealth in the United States”. In: *Journal of Economic Interaction and Coordination* 13.3 (2018), pp. 641–656 (cit. on pp. 3, 96, 111).

[125] Philip Vermeulen. “How fat is the top tail of the wealth distribution?” In: *Review of Income and Wealth* 64.2 (2018), pp. 357–387 (cit. on pp. 1, 95).

[126] W.A. Waugh. “Modes of growth of counting processes with increasing arrival rates”. In: *Journal of Applied Probability* 11.2 (1974), pp. 237–247 (cit. on p. 70).

[127] W.A. Waugh. “Uses of the sojourn time series for Markovian birth process”. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Probability Theory*. Vol. 6. University of California Press. 1972, pp. 501–515 (cit. on p. 70).

[128] Ward Whitt. “Proofs of the martingale FCLT”. In: *Probability Surveys* 4 (2007), pp. 268–302 (cit. on pp. 58, 61).

[129] Herman O.A. Wold and Peter Whittle. “A model explaining the Pareto distribution of wealth”. In: *Econometrica* 25.4 (1957), pp. 591–595 (cit. on p. 96).

[130] Victor M. Yakovenko and J. Barkley Rosser Jr. “Colloquium: Statistical mechanics of money, wealth, and income”. In: *Reviews of modern physics* 81.4 (2009), p. 1703 (cit. on p. 96).

[131] George U. Yule. “A mathematical Theory of evaluation, based on conclusions of Dr. J. Willis”. In: *Philosophical Transactions* (1924), pp. 213–221 (cit. on p. 69).

[132] Tong Zhu. “Nonlinear Pólya urn models and self-organizing processes”. PhD thesis. University of Pennsylvania, Philadelphia, 2009 (cit. on pp. 11, 13, 69, 74, 81, 88, 163, 168).

Appendix A

Exponentially Decreasing Feedback

Based on an example, this supplemental chapter discusses the long-time limits of a Pólya urn with exponentially decreasing feedback, since this case is not covered by the results of Section 3.2.

Example A.1. Let $A = 2$ and $F_i(k) = \alpha_i e^{-\beta_i k}, \alpha_i, \beta_i > 0, i = 1, 2$. As explained in detail in Section 4.1, we can write

$$\chi_1(n) = \chi_1(0) + H_1(n) + M_1(n) \quad \text{for } n \geq 0 \ ,$$

where $(M_1(n))_{n \in \mathbb{N}_0}$ is an almost surely convergent martingale and

$$H_1(n) \coloneqq \sum_{k=0}^{n-1} \frac{G_1(N+k, \chi_1(k))}{N+k+1}$$

is predictable with $G_1(k, x) \coloneqq p_1(k, (x, 1-x)) - x$, $x \in (0, 1)$ given by centered transition probabilities (2.2). In the case of exponentially decreasing feedback, we have the following convergence:

$$G_1(k, x) \xrightarrow{k \to \infty} G_1(x) \coloneqq \begin{cases} 1 - x, & \text{if } x\beta_1 < (1-x)\beta_2 \\ \frac{\alpha_1}{\alpha_1 + \alpha_2} - x, & \text{if } x\beta_1 = (1-x)\beta_2 \\ -x, & \text{otherwise} \end{cases} \tag{A.1}$$

The convergence is locally uniform in $(0, 1)$ apart from the point $x = x_0 \coloneqq \frac{\beta_2}{\beta_1 + \beta_2}$. Take $\epsilon > 0$. For large enough k, $G_1(k, (\cdot))$ is sufficiently close to G_1 outside an ϵ-neighborhood of x_0. If for a large n, $|\chi_1(n) - x_0| > \epsilon$, then the process $(\chi_1(n))_n$ enters the ϵ-neighborhood of x_0 in finite time because of the convergence of the martingale. As the same holds for $\epsilon/2$ instead of ϵ, we get that the process leaves this ϵ-neighborhood only finitely often. This yields

$$\chi_1(n) \xrightarrow{n \to \infty} x_0 \quad \text{almost surely.}$$

Thus, the limit is not only independent of the initial market shares, but also of the fitness-parameters α_i (in contrast to polynomially decreasing feedback). Note that these findings are consistent with Corollary 3.26, i.e. (3.20) still holds. Because of the independence property in the exponential embedding in Section 2.1, this can easily be extended to general A. For different (at least) exponentially decreasing feedback, we basically only need a convergence as in (A.1) for an analogous result.

Remarkably, Example A.1 reveals the following behavioural difference between exponentially decreasing and polynomial feedback. Suppose that there are agents i, j such that

$$\lim_{k\to\infty} \frac{F_i(k)}{F_j(k)} = 0.$$

Then agent i is marginalized, i.e. $\lim_{n\to\infty} \chi_i(n) = 0$, if F_i satisfies (3.21), in particular if $F_i(k) = \alpha_i k^{\beta_i}$ for $\beta_i < 1$. On the other hand, for exponentially decreasing feedback like in Example A.1, we might still have $\lim_{n\to\infty} \chi_i(n) > 0$.

Appendix B

Functional Limit Theorems with Non-Linear Time Scale

From a stochastical point of view, the first steps of a generalized Pólya urn are of special interest because the randomness plays a significant role. In the later stages of the process, the market shares and thus the probability of winning in a certain step remain almost invariant, such that the sequence of winners $(X(n+1)-X(n))_n$ is almost independent and identically distributed for large n. Even in the Central Limit Theorem 4.7 the limiting process M becomes virtually constant for large t. In order to particularly focus on the early stages of the process, we analyse the process $N^{1-\frac{\beta}{2}}\left(\chi^{(N)}(\lfloor N^\beta t\rfloor)\right)_{t>0}$ for large initial market size N and $\beta\in(0,1)$. Recall the Doob decomposition (4.4) and the notation from Section 4.1.

Theorem B.1. *Suppose that the assumptions of Theorem 4.7 are fulfilled and denote by $(B_t)_{t\geq 0}$ a standard Brownian motion. Then, we have weak convergence to a Brownian motion for any $\beta\in(0,1)$:*

$$N^{1-\frac{\beta}{2}}\left(M_1^{(N)}(\lfloor N^\beta t\rfloor)\right)_{t\geq 0}\xrightarrow{N\to\infty}\sqrt{p_1(\chi(0))(1-p_1(\chi(0)))}\,(B_t)_{t\geq 0}\quad \textit{weakly on } \mathbb{D}([0,\infty),\mathbb{R})\,.$$

Proof. We will only sketch the proof as it is quite analogous to the proof of Theorem 4.7. We use the tightness given by Lemma 4.8 and assume that the sequence $N^{1-\frac{\beta}{2}}\left(M_1^{(N)}(\lfloor N^\beta(\cdot)\rfloor)\right)_N$ converges to a process $\hat{M}_1$. Then we take a smooth test-function $f\colon\mathbb{R}\to\mathbb{R}$ with compact support and consider the martingales

$$f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(\lfloor N^\beta t\rfloor)\right)-f(0)-\sum_{k=0}^{\lfloor N^\beta t\rfloor-1}\mathbb{E}\left[f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k+1)\right)-f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)\big|\mathcal{F}_k^{(N)}\right].$$

Then we know that $f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(\lfloor N^\beta t\rfloor)\right)$ converges to $f\left(\hat{M}_1(t)\right)$ and via Lemma 4.9 we get:

$$\sum_{k=0}^{\lfloor N^\beta t\rfloor-1} \mathbb{E}\left[f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k+1)\right)-f\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)|\mathcal{F}_k^{(N)}\right]$$
$$=\sum_{k=0}^{\lfloor N^\beta t\rfloor-1}\frac{N^{2-\beta}}{2(N+k+1)^2}f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}(k)\right)p_1(N+k,\chi^{(N)}(k))\left(1-p_1(N+k,\chi^{(N)}(k))\right)$$
$$=\sum_{k=0}^{\lfloor N^\beta t\rfloor-1}\frac{f''\left(N^{1-\frac{\beta}{2}}M_1^{(N)}\left(N^\beta\frac{k}{N^\beta}\right)\right)}{2N^\beta(1+\frac{k}{N}+\frac{1}{N})^2}p_1\left(N+k,Z^{(N)}\left(\frac{k}{N}\right)\right)\left(1-p_1\left(N+k,Z^{(N)}\left(\frac{k}{N}\right)\right)\right)$$
$$\xrightarrow{N\to\infty}\int_0^t\frac{f''(\hat{M}_1(s))}{2}p_1(Z(0))\left(1-p_1(Z(0)\right)ds\ ,$$

where we have used $\beta<1$ and $k/N\to 0$ in the last step. This implies that $\hat{M}_1$ is a Markov process with generator $p_1(Z(0))\left(1-p_1(Z(0)\right)\frac{f''}{2}$. Hence, $\hat{M}_1$ is the desired Brownian motion. □

Note that Theorem B.1 is consistent with Theorem 4.7 for small t. This limiting Brownian motion can be understood as a consequence of Donsker's invariance principle, since the shares do barely change at the beginning of the process for large initial values. Again, a straight forward extension to higher dimensions is possible.

Theorem B.2. *Suppose that the assumptions of Theorem 4.7 are fulfilled and let $\beta\in(0,1)$. Then the sequence of processes $N^{1-\frac{\beta}{2}}\left(M^{(N)}(\lfloor N^\beta t\rfloor)\right)_{t\geq 0}$ converges for $N\to\infty$ to a time-homogeneous Markov process with generator*

$$\hat{L}f(x):=\frac{1}{2}\sum_{\substack{i,j=1\\ i<j}}^{A}p_i(Z(0))p_j(Z(0))\left(\frac{\partial}{\partial x_i}-\frac{\partial}{\partial x_j}\right)^2 f(x),\quad x\in\mathbb{R}^A$$

weakly on $\mathbb{D}([0,\infty),T\Delta_{A-1})$.

As in Subsection 4.2.1, the limit process can be interpreted as independent exchanges of mass between pairs of agents according to a Brownian motion.

We already know from Theorem 4.1 that $\chi^{(N)}(\lfloor N^\beta t\rfloor)$ converges to $\chi(0)$ for $N\to\infty$, when $\beta<1$. Moreover, Theorem B.2 states that the process $\left(M^{(N)}(\lfloor N^\beta t\rfloor)\right)_{t\geq 0}$ converges to zero at rate $N^{\frac{\beta}{2}-1}$. In addition, it follows from

$$N^{1-\beta}H^{(N)}(\lfloor N^\beta t\rfloor)=N^{1-\beta}\sum_{k=0}^{\lfloor N^\beta t\rfloor-1}\frac{G(N+k,\chi^{(N)}(k))}{N+k+1}$$
$$=\sum_{k=0}^{\lfloor N^\beta t\rfloor-1}\frac{1}{N^\beta}\cdot\frac{G(N+k,\chi^{(N)}(N^\beta\cdot\frac{k}{N^\beta}))}{1+\frac{k}{N}+\frac{1}{N}}\xrightarrow{N\to\infty}\int_0^t G(\chi(0))du=G(\chi(0))t,$$

that $N^{1-\beta}\left(H^{(N)}(\lfloor N^\beta t\rfloor)\right)_{t\geq 0}$ converges to $(G(\chi(0))t)_{t\geq 0}$, which immediately implies the following law of large numbers.

Corollary B.3. *Under the assumptions of Theorem B.1 we have*

$$N^{1-\beta}\left(\chi^{(N)}(\lfloor N^\beta t\rfloor)-\chi(0)\right)_{t\geq 0}\xrightarrow{N\to\infty}(G(\chi(0))t)_{t\geq 0}\quad \textit{weakly on } \mathbb{D}([0,\infty),\mathbb{R}^A)\ .$$

Combining these results for an analysis of the deviations of $\chi^{(N)}(\lfloor N^\beta t\rfloor)$ requires further distinction of β as specified in the following functional CLT.

Corollary B.4. *Let $\beta\in(0,1)$ and $\gamma>0$ as specified below. Suppose that*

$$\lim_{k\to\infty}k^\gamma\sup_{x\in\Delta_{A-1}}\|G(k,x)-G(x)\|=0.$$

Moreover, let G be continuously differentiable. Then

$$N^\gamma\left(N^{1-\beta}\left(\chi^{(N)}(\lfloor N^\beta t\rfloor)-\chi(0)\right)-G(\chi(0))t\right)_{t\geq 0}\xrightarrow{N\to\infty}(\hat{Z}(t))_{t\geq 0}\quad \textit{weakly on } \mathbb{D}([0,\infty),\mathbb{R}^A)\ ,$$

where the limiting process $\hat{Z}$ is defined as follows:

1. *For $\beta>\frac{2}{3}$ set $\gamma=1-\beta$. Then:*

$$\hat{Z}(t)=\frac{1}{2}DG(\chi(0))\cdot G(\chi(0))t^2.$$

2. *For $\beta=\frac{2}{3}$ set $\gamma=\frac{1}{3}$. Then*

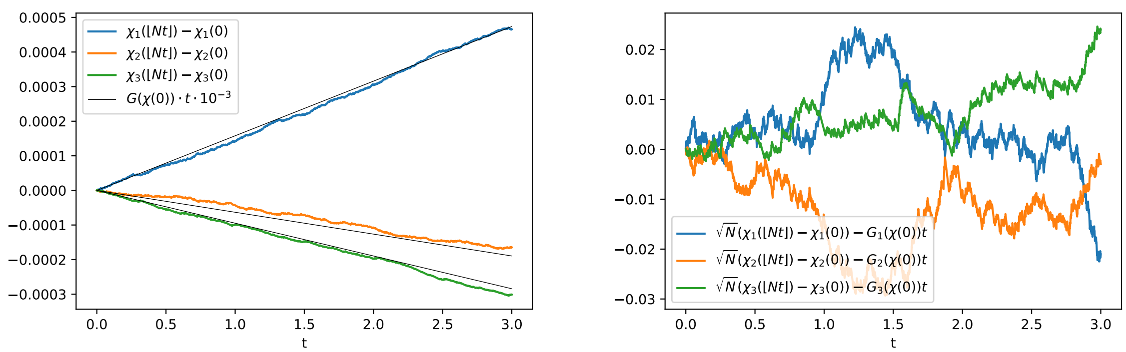
$$\hat{Z}(t)=\frac{1}{2}DG(\chi(0))\cdot G(\chi(0))t^2+\hat{M}(t),$$

where $\hat{M}$ is the limiting process from Theorem B.2.

3. *For $\beta<\frac{2}{3}$ set $\gamma=\frac{\beta}{2}$. Then $\hat{Z}=\hat{M}$.*

Proof. We only sketch the proof as it is widely analogous to the proof of Theorem 4.12. Again, we use Doob's decomposition (4.4) and rephrase as follows:

$$\begin{aligned}
&N^\gamma\left(N^{1-\beta}\left(\chi^{(N)}(\lfloor N^\beta t\rfloor)-\chi(0)\right)-G(\chi(0))t\right)\\
&=N^\gamma\left(N^{1-\beta}H^{(N)}(\lfloor N^\beta t\rfloor)-G(\chi(0))t\right)+N^{1-\beta+\gamma}M^{(N)}(\lfloor N^\beta t\rfloor)\\
&=N^\gamma\int_0^t\left(G(\chi^{(N)}(N^\beta u))-G(\chi(0))\right)du+o(1)+N^{1-\beta+\gamma}M^{(N)}(\lfloor N^\beta t\rfloor)\\
&=N^\gamma\int_0^t DG(\chi(0))\cdot\left(\chi^{(N)}(N^\beta u)-\chi(0)\right)du+o(1)+N^{1-\beta+\gamma}M^{(N)}(\lfloor N^\beta t\rfloor)\\
&=N^{\gamma+\beta-1}DG(\chi(0))\cdot\int_0^t N^{1-\beta}\left(\chi^{(N)}(N^\beta u)-\chi(0)\right)du+o(1)+N^{1-\beta+\gamma}M^{(N)}(\lfloor N^\beta t\rfloor)
\end{aligned}$$



(a) (b)

Figure B.1: Simulation of the processes $\chi^{(N)}(\lfloor N^\beta t \rfloor) - \chi(0)$ (left) and $N^{1-\beta}\left(\chi^{(N)}(\lfloor N^\beta t \rfloor) - \chi(0)\right) - G(\chi(0))t$ (right) for $\beta = \frac{1}{2}$ and $N = 10^6$. We took $A = 3$, $F_1(k) = F_2(k) = F_3(k) = k^2$ and $\chi(0) = (0.5, 0.3, 0.2)$.

Finally, the claims follow via Theorem B.2 and Corollary B.3. □

The assumptions of Theorem $B.4$ are satisfied e.g. for $F_i(k) = \alpha_i k^\beta$. To sum up, in the limit $N \to \infty$ the process $\chi^{(N)}(\lfloor N^\beta t \rfloor)$ stays at $\chi(0)$ for all time t. After scaling, Corollary B.3 reveals a linear drift into direction $G(\chi(0))$. The fluctuations around this linear drift can itself be described by a random SDE for $\beta \leq \frac{2}{3}$ and by a deterministic ODE for $\beta > \frac{2}{3}$, since second order terms dominate the randomness for $\beta > \frac{2}{3}$. These findings are illustrated by Figure B.1.

Appendix C

Supplemental Material for Chapter 6

C.1 A law of large numbers for the dynamics

In complete analogy to the $r = 0$ case in Section 4.1, it is possible to describe the extended process $X(n)$ defined in Section 6.2 in the limit for large initial values via a functional LLN, where the limiting process is deterministic and given by the solution of an ODE.

The key for the proof will be once again the Doob decomposition (6.4), where we can find separate limits for the predictable part H and the martingale part M. In order to emphasize the dependence of the initial market size, we will write $\chi^{(N)} := \chi$, $X^{(N)} := X$ and $H^{(N)} := H$, $M^{(N)} := M$ in the following, where $N := X_1^{(N)}(0) + \ldots + X_A(N)(0)$. We keep the initial wealth distribution $\chi^{(N)}(0) = \chi(0) \in \Delta^o_{A-1}$ fixed. For simplicity, we assume homogeneous feedback $F_i(k) = \alpha_i k^\beta, \alpha_i > 0, \beta \in \mathbb{R}$ for this section, such that $G(n, x) = G(x)$ and $p(n, x) = p(x)$ do not depend on the market size n. Assuming sufficiently fast uniform convergence of $G(n, (\cdot))$ towards a field G, the results can be extended to non-homogeneous feedback like in Section 4.1. We now formulate our LLN.

Theorem C.1. *In the situation of Chapter 6 with $F_i(k) = \alpha_i k^\beta, \alpha_i > 0, \beta \in \mathbb{R}$, denote by $\tilde{Z} \colon [0, \infty) \to \Delta_{A-1}$ the solution of the ODE*

$$\frac{d}{dt}\tilde{Z}(t) = \frac{G(Z(t))}{1+t} \quad \textit{with} \quad \tilde{Z}(0) = \chi(0) \in \Delta^o_{A-1} \tag{C.1}$$

and define the sequence of processes in Δ_{A-1}

$$\left(\tilde{Z}^{(N)}\right)_N := \left(\tilde{Z}^{(N)}(t) \colon t \geq 0\right)_N := \left(\chi^{(N)}(\lfloor Nt \rfloor) \colon t \geq 0\right)_N .$$

Then: $\tilde{Z}^{(N)}$ converges to $\tilde{Z}$ weakly on the Skorochod space $\mathbb{D}([0, \infty), \Delta_{A-1})$.

Note that G is locally Lipschitz, such that $\tilde{Z}$ is unique and well-defined. An important implication of our LLN is the following: It is possible that $\tilde{Z}$ does converge towards a saddle point of G, although our process $\chi^{(N)}$ cannot converge to saddle points due to noise. Hence, the process $\chi^{(N)}$ can be stuck in saddle points for quite a long time, i.e. it takes more than $O(N)$ time to escape from a saddle point.

The proof for the $r = 0$ case in Section 4.1 can be transferred one-to-one to $r > 0$, since all crucial properties do also hold in the general case. These are in particular:

1. $\|\tilde{Z}^{(N)}(t) - \tilde{Z}^{(N)}(s)\|_\infty \leq \frac{N|t-s|+1}{N}$ holds for $t, s \geq 0$ implying tightness of the sequence $\tilde{Z}^{(N)}$.

2. The the increments $\xi_i^{(N)}(n)$ are centered and uniformly bounded. Moreover, $\xi_i^{(N)}(n)$ and $\xi_j^{(N)}(m)$ are uncorrelated for $n \neq m$. Hence, Doob's inequality yields that $\left(M^{(N)}(\lfloor Nt \rfloor) \colon t \geq 0\right)$ converges to zero for $N \to \infty$ (weakly on $\mathbb{D}([0,\infty), \Delta_{A-1})$).

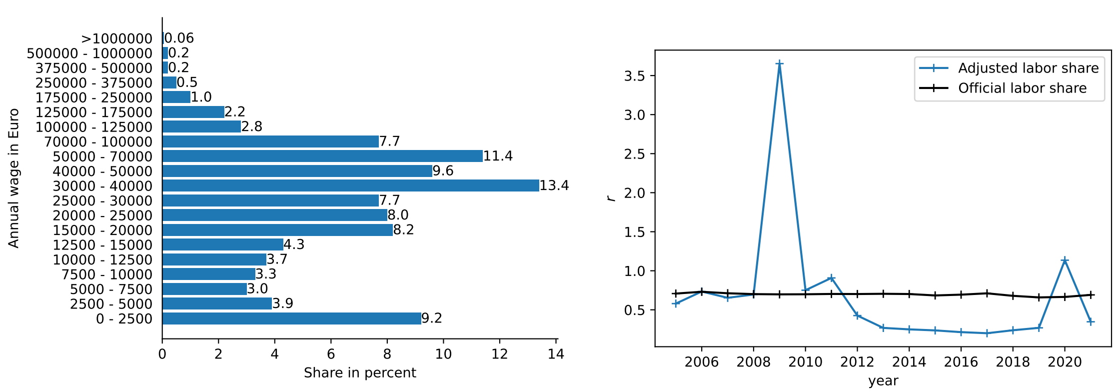
3. Riemann approximation of the integral yields

$$H^{(N)}(\lfloor Nt \rfloor) = \sum_{k=0}^{\lfloor Nt \rfloor - 1} \frac{1}{N} \cdot \frac{G\left(\chi^{(N)}(N \cdot \frac{k}{N})\right)}{1 + \frac{k}{N} + \frac{1}{N}} \xrightarrow{N \to \infty} \int_0^t \frac{G(\tilde{Z}(t))}{1+u} du = \tilde{Z}(t) - \tilde{Z}(0),$$

which completes the proof.

Thus, the process behaves almost deterministic for large initial values.

C.2 Supplementary figures



(a) (b)

Figure C.1: (a) shows the distribution of annual net wages in Germany 2018 per taxpayer based on [31]. (b) presents empirical values of the labor share r for several years in Germany computed according to formula (6.11) using data from [111] and [32] compared to the official labor share from [33].

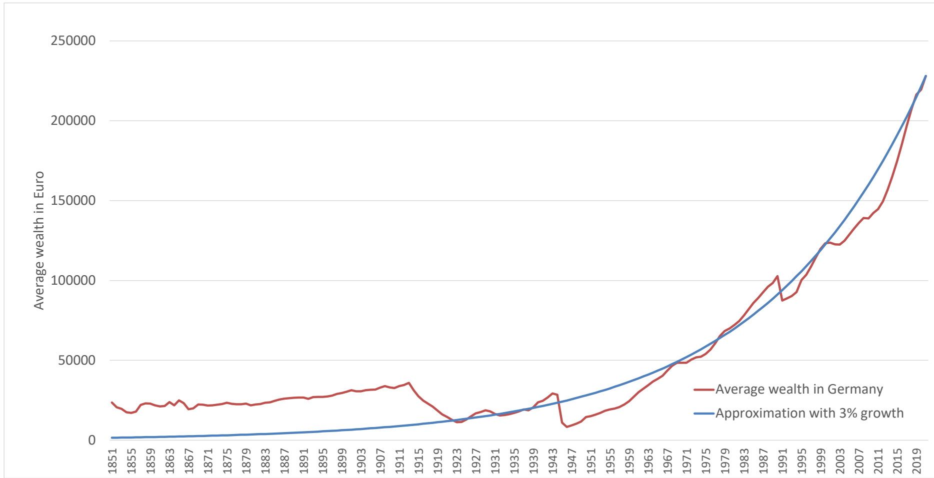
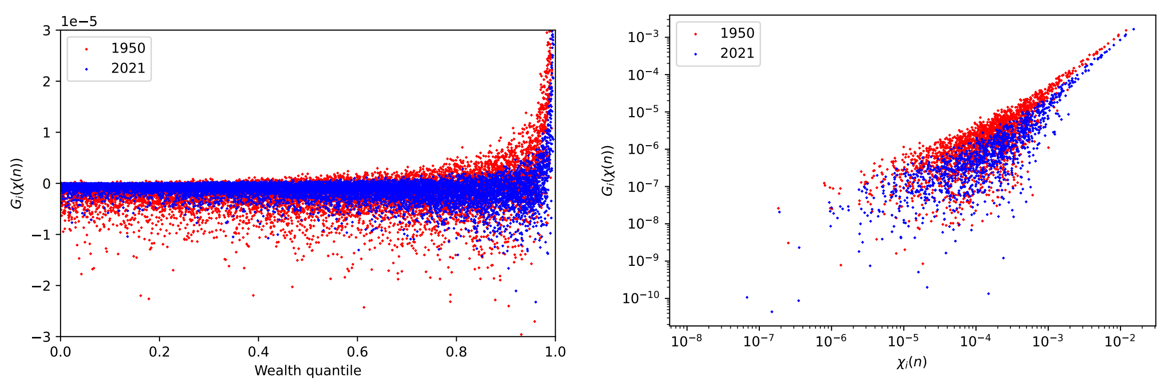


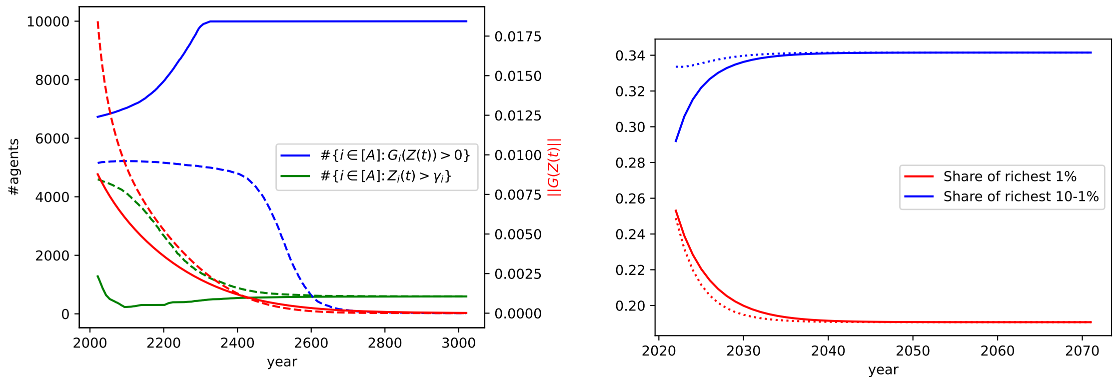
Figure C.2: Timeline of average personal wealth in Germany according to [111] compared to constant growth at rate 3%. This justifies the parameter choice $\mu = 0.03$ in the time scaling (6.12) in Section 6.4.5.



(a) all points

(b) only positive values of $G_i(\chi(n))$

Figure C.3: Scatter plot of $G_i(\chi(n))$ against wealth quantiles (a) and against $\chi_i(n)$ (b), for the years 1950 and 2021 in our simulation with symmetric initial condition. Years were assigned via (6.12) assuming a constant growth rate of $\mu = 0.03$ per year. The entries of the field G slowly tend to zero asymptotically and this happens faster for poorer agents (cf. Section 6.4.5).



(a) Figure 6.13 (b) Figure 6.14

Figure C.4: Analogoue figures to Figure 6.13 and Figure 6.14 supposing $r = 0.5$.

C.3 Heuristics on the stability of fixed points

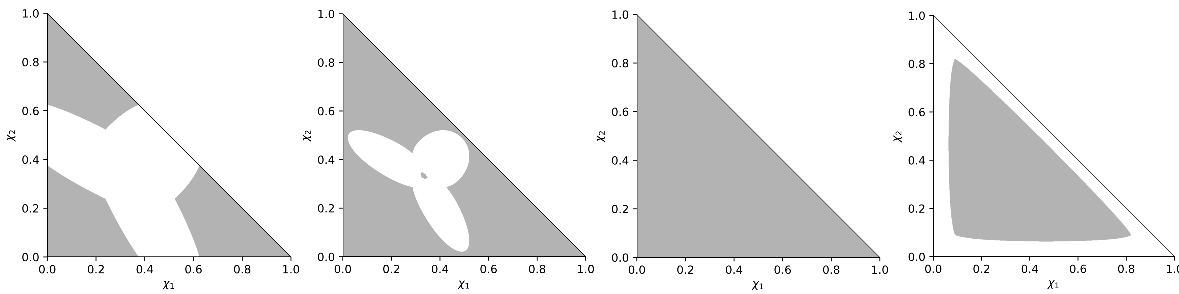
In Subsection 6.4.4, we faced the challenge to decide whether a given fixed point of the field G is stable. Formally, we would have to check negative definiteness of the Hessian of the Lyapunov function (6.6) using e.g. Lanczos' algorithm. Since this is numerically difficult for large A, we are content with some heuristics. These arguments will also grant some further insight into the possible positions of stable fixed points. We recall the definition of the field G in (6.2) and (6.5). We assume homogeneous feedback $F_i(k) = k^\beta$, $\beta \in \mathbb{R}$ in the following.

A fixed point $x \in \Delta_{A-1}$ of G is stable if any infinitesimal exchange of mass between two agents has an inverse effect on G, i.e.

$$\frac{\partial G_i(x)}{\partial x_i} - \frac{\partial G_i(x)}{\partial x_j} < 0 \quad \text{for all } j \neq i\,. \tag{C.2}$$

If $\frac{\partial G_i(x)}{\partial x_i} - \frac{\partial G_i(x)}{\partial x_j} > 0$ holds for one pair $i \neq j$, then any increase of x_i at the expense of agent j would even be reinforced by the field G, such that x is unstable. The partial derivatives can easily be computed:

$$\frac{\partial G_i(x)}{\partial x_i} = (1-r)\frac{\beta x_i^{\beta-1}\sum_k x_k^\beta - \beta x_i^\beta x_i^{\beta-1}}{\left(\sum_k x_k^\beta\right)^2} - 1 = (1-r)\frac{\beta}{x_i}\left(p_i(x) - p_i(x)^2\right) - 1$$



(a) $r = 0.4$, $\beta = 2$ (b) $r = 0.51$, $\beta = 2$ (c) $r \geq 0.6$, $\beta = 2$ (d) $r = 0.1$, $\beta = 0.5$

Figure C.5: The set P for $A = 3$ and different r, β.

And for $j \neq i$:

$$\frac{\partial G_i(x)}{\partial x_j} = -(1-r)\frac{x_i^\beta}{\left(\sum_k x_k^\beta\right)^2} x_j^{\beta-1}\beta = -(1-r)\frac{\beta}{x_i} p_i(x)^2 \left(\frac{x_j}{x_i}\right)^{\beta-1}$$

Thus:

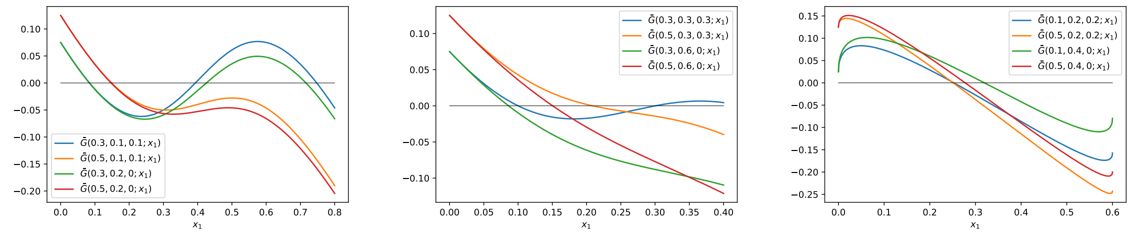
$$\begin{aligned}\frac{\partial G_i(x)}{\partial x_i} - \frac{\partial G_i(x)}{\partial x_j} &= (1-r)\frac{\beta}{x_i}\left(p_i(x) + p_i(x)^2\left(\left(\frac{x_j}{x_i}\right)^{\beta-1} - 1\right)\right) - 1 \\ &= (1-r)\beta\frac{p_i(x)}{x_i}\left(1 + p_i(x)\left(\left(\frac{x_j}{x_i}\right)^{\beta-1} - 1\right)\right) - 1\end{aligned}$$

First, note that this condition does not depend on γ, but the position of fixed points does so of course. Moreover, it suffices to only take the richest agent for j in (C.2) since $\frac{\partial G_i(x)}{\partial x_j}$ is monotone in x_j. Hence, we only have to check $A-1$ inequalities, such that the criterion is numerically fast.

If $\beta > 1$, then obviously (C.2) holds for $x = e^{(k)}$ for all $k \in [A]$. Since (C.2) does continuously depend on x, all fixed points, which are close to a corner of the simplex, are stable. On the other hand, the point $x = \left(\frac{1}{\#S}\mathbb{1}_{i \in S}\right)_{i \in [A]}$, where the total wealth is shared among agents $S \subset [A]$ with $\#S > 1$, fulfills (C.2) only if

$$r > \frac{\beta - 1}{\beta}.$$

Hence, the set $P := \{x \in \Delta_{A-1} \colon (C.2) \text{ holds}\}$ provides the region in Δ_{A-1} where stable fixed points can exist. It is increasing in r, does always contain the corners of the simplex and contains the middle point of the simplex for $r > \frac{\beta-1}{\beta}$. Figure C.5 illustrates this set P. Therefore it is plausible that for symmetric wage $\gamma = \left(\frac{1}{A}, \ldots, \frac{1}{A}\right)$ the critical labor share is $r_c = \frac{\beta-1}{\beta}$ which is consistent with our considerations of the $A = 2$ case in Section 6.3. Remarkably, this expression does not depend on A.



(a) $\beta = 2$, $x_3 + x_4 = 0.2$ (b) $\beta = 2$, $x_3 + x_4 = 0.6$ (c) $\beta = 0.5$, $x_3 + x_4 = 0.4$

Figure C.6: The mapping $\bar{G}(r, x_3, x_4; x_1) \coloneqq (1-r)\frac{x_1^\beta}{x_1^\beta + (1-x_1-x_3-x_4)^\beta + x_3^\beta + x_4} + 0.25r - x_1$ for various parameterizations.

For $\beta = 1$, the non-linear part G_0 (6.7) vanishes and the total field is $G_i(x) = r(\gamma_i - x_i)$. Therefore we have $\frac{\partial G_i(x)}{\partial x_i} - \frac{\partial G_i(x)}{\partial x_j} = -r \leq 0$, the region $P = \Delta_{A-1}$ and the unique fixed point $x = \gamma$ is stable. For $\beta < 1$, the symmetric point $x = \left(\frac{1}{A}, \ldots, \frac{1}{A}\right)_{i \in [A]}$ fulfills (C.2) for all $r \geq 0$, but P does not contain any boundary point of the simplex Δ_{A-1} (Figure C.5 (d)). Thus, any stable point must be located in the interior of the simplex.

Finally, we add another heuristic to explain, why stable fixed points are either close to the vertices of the simplex or there is only one stable fixed point. For that, we take two agents $i \neq j$ and fix the shares of the others. Define $c \coloneqq \sum_{\substack{k \in [A] \\ k \neq i,\, k \neq j}} x_k$ and $d \coloneqq \sum_{\substack{k \in [A] \\ k \neq i,\, k \neq j}} x_k^\beta$. Then we consider the field G while distributing the remaining $1 - c$ share:

$$x_i \mapsto (1-r)\frac{x_i^\beta}{x_i^\beta + (1-c-x_i)^\beta + d} + r\gamma_i - x_i$$

Via plotting (Figure C.6), one can easily see that for $\beta \leq 1$ or c, r large enough, this function has only one fixed point. Otherwise, there are three fixed points, but the middle one is unstable. Hence, stable fixed points are either unique and in the middle of the simplex ($r > r_c'$), or random and close to the corners of the simplex ($r < r_c'$). In the latter case, there are unstable fixed points and saddle points away from the corners.

Appendix D

Systems with Many Agents

In some applications of the generalized Pólya urn model, the number of agents is typically large, e.g. when then wealth distribution within a nation is modelled. Although most of our results hold for general finite A, we now take a closer look at the limiting case $A \to \infty$.

D.1 General behaviour for $A \to \infty$

We fix a sequence $F_1, F_2, \dots$ of feedback functions and a sequence $X_1(0), X_2(0), \dots \in \mathbb{N}$ of initial values. Denote by $(X^{(A)}(n))_{n \geq 0}$ a generalized Pólya urn with A agents and feedback functions $F_1, \dots, F_A$ and initial values $X^{(A)}(0) = (X_1(0), \dots, X_A(0))$ as defined in Section 2.1. In addition, define

$$X_i^{(A)}(\infty) := \lim_{n\to\infty} X_i^{(A)}(n) \in \mathbb{N} \cup \{\infty\}.$$

The approach is based on the exponential embedding described in Section 2.1. So, let $\Xi_1, \Xi_2, \dots$ be a sequence of independent birth processes with initial values $X_1(0), X_2(0), \dots$ and rate functions $F_1, F_2, \dots$. We denote by $\tau_i(k)$ the sojourn times of Ξ_i and by $T_i := \sum_{k=X_i(0)}^{\infty} \tau_i(k) \in (0, \infty]$ the explosion time of Ξ_i. In the following, $\mathbb{P}$ denotes the law of the sequence of birth processes and for each $A \geq 1$ the process $(X^{(A)}(n))_{n \geq 0}$ is then defined as in Section 2.1.

First, we consider the case when infinitely many F_i satisfy the monopoly condition (M). Theorem 5.10 describes the tail distribution of a loser among those agents. Although the tail decay rate does not depend on A, the constant prefactor vanishes for $A \to \infty$. Indeed, a typical agent does not even win a single step in large systems.

Theorem D.1. *Assume that there is an infinite subset $B \subset \mathbb{N}$, such that*

$$\sum_{k=0}^{\infty} \max_{i \in B} \frac{1}{F_i(k)} < \infty. \tag{D.1}$$

Then:

1. *The convergence*
$$\lim_{A \to \infty} \mathbb{P}\left(X_1^{(A)}(\infty) = X_1(0)\right) = 1$$
holds, i.e. $X_1^{(A)}(\infty) \xrightarrow{A \to \infty} X_1(0)$ *in distribution.*

2. *If in addition* $\limsup_{i \to \infty} F_i(X_i(0)) < \infty$, *then*
$$\frac{1}{A} \#\{i \in [A] \colon X_i^{(A)}(\infty) = X_i(0)\} \xrightarrow{A \to \infty} 1 \quad \textit{in distribution.}$$

3. *If* $\liminf_{i \to \infty} F_i(X_i(0)) > 0$ *and* $\lim_{i \to \infty} \frac{1}{i} F_i(X_i(0) + 1) = 0$, *then*
$$\#\{i \in [A] \colon X_i^{(A)}(\infty) \neq X_i(0)\} \xrightarrow{A \to \infty} \infty \quad \textit{in distribution.}$$

In other words: In the limit $A \to \infty$, almost all agents do not win a single step. Nevertheless, the absolute number of agents winning at least one step is unbounded. Note that (D.1) implies that infinitely many agents fulfill (M). Obviously, (D.1) is satisfied when infinitely many F_i are equal and satisfy (M), a typical counterexample would be $F_i(k) = k^{1+\frac{1}{i}}$. The assumption (D.1) can be replaced by any other condition, that ensures that $\min_{i=1,\dots,A} T_i \xrightarrow{A \to \infty} 0$ in distribution.

Lemma D.2. *Assume that (D.1) holds. Then:*

$$\min_{i=1,\dots,A} T_i \xrightarrow{A \to \infty} 0 \quad \textit{in distribution.} \tag{D.2}$$

Proof. We fix $\epsilon > 0$ and show that $\mathbb{P}\left(\min_{i=1,\dots,A} T_i > \epsilon\right) \xrightarrow{A \to \infty} 0$. For that, we take a sequence $(\tilde{\tau}(k))_k$ of independent random variables, where $\tilde{\tau}_i(k)$ is exponentially distributed with expectation $\max_{i \in B} \frac{1}{F_i(k)}$. Then $\tilde{T} \coloneqq \sum_{k=1}^{\infty} \tilde{\tau}(k)$ is almost surely finite due to assumption (D.1). Moreover, $\tilde{T}$ is stochastically bigger than all T_i with $i \in B$. Using the independence of the T_i, we can now estimate

$$\mathbb{P}\left(\min_{i=1,\dots,A} T_i > \epsilon\right) = \prod_{i=1}^{A} \mathbb{P}(T_i > \epsilon) \leq \prod_{\substack{i \in B \\ i \leq A}} \mathbb{P}(T_i > \epsilon) \leq \prod_{\substack{i \in B \\ i \leq A}} \mathbb{P}(\tilde{T} > \epsilon) \xrightarrow{A \to \infty} 0$$

since B is infinite and $\mathbb{P}(\tilde{T} > \epsilon) < 1$ as a sum of exponential random variables. ☐

Now, the independence property of the exponential embedding allows a canonical coupling of the processes $X^{(A)}$, which helps to prove Theorem D.1.

Proof of Theorem D.1. All independent random variables $\tau_i(k), i \geq 1, k \geq 1$ are defined on the same probability space with distribution $\mathbb{P}$. The sequence $(\min_{i=2,\ldots,A} T_i)_A$ is decreasing in A and hence almost surely convergent with limit zero (due to Lemma D.2). Now, fix a realisation. Then $\tau_1(X_1(0)) > \min_{i=2,\ldots,A} T_i$ holds for A large enough, i.e. $X_1^{(A)}(\infty) = X_1(0)$. Thus, $X_1^{(A)}(\infty)$ converges to $X_1(0)$ almost surely for $A \to \infty$, which implies part 1 of Theorem D.1.

For part 2, we observe that

$$T_{2:A} := \min\left\{T_i \colon i \in [A] \text{ and } T_i \neq \min_{j=1,\ldots,A} T_j\right\} \xrightarrow{A\to\infty} 0 \quad \text{almost surely}$$

holds, too. Due to dominated convergence, we even have $\mathbb{E}e^{-cT_{2:A}} \xrightarrow{A\to\infty} 1$ for any constant $c > 0$. In addition, define $T_{[A]\setminus\{i\}} := \min_{j\in[A]\setminus\{i\}} T_j$, which obviously satisfies $T_{[A]\setminus\{i\}} \leq T_{2:A} \to 0$. Now we are prepared for the final calculation:

$$\begin{aligned}
&\mathbb{E}\left(\frac{\#\{i \in [A] \colon X_i^{(A)}(\infty) = X_i(0)\}}{A}\right) = \mathbb{E}\left(\frac{1}{A}\sum_{i=1}^{A} \mathbb{1}_{X_i^{(A)}(\infty)=X_i(0)}\right) \\
&= \frac{1}{A}\sum_{i=1}^{A} \mathbb{P}\left(X_i^{(A)}(\infty) = X_i(0)\right) = \frac{1}{A}\sum_{i=1}^{A} \mathbb{E}\left[\mathbb{P}\left(X_i^{(A)}(\infty) = X_i(0) \,\middle|\, T_{[A]\setminus\{i\}}\right)\right] \\
&= \frac{1}{A}\sum_{i=1}^{A} \mathbb{E}\left[\mathbb{P}\left(\tau_i(X_i(0)) > T_{[A]\setminus\{i\}} \,\middle|\, T_{[A]\setminus\{i\}}\right)\right] = \frac{1}{A}\sum_{i=1}^{A} \mathbb{E}\exp\left(-F_i(X_i(0))T_{[A]\setminus\{i\}}\right) \\
&\geq \frac{1}{A}\sum_{i=1}^{A} \mathbb{E}\exp\left(-F_i(X_i(0))T_{2:A}\right) \geq \mathbb{E}\exp\left(-const.T_{2:A}\right) \xrightarrow{A\to\infty} 1
\end{aligned}$$

This is sufficient for 2. since $\frac{1}{A}\#\{i \in [A] \colon X_i^{(A)}(\infty) = X_i(0)\} \leq 1$.

For part 3, we take any $k \in \mathbb{N}$ and show, that the probability, that the first k steps are won by k different agents, converges to one for $A \to \infty$:

$$\begin{aligned}
&\mathbb{P}\left(\#\{i \in [A] \colon X_i^{(A)}(\infty) \neq X_i(0)\} \geq k\right) \\
&\quad \geq \mathbb{P}\left(X(n) - X(n-1) \neq X(m) - X(m-1) \text{ for all } n, m = 1, \ldots, k,\ n \neq m\right) \\
&\quad \geq \prod_{n=1}^{k}\left(1 - \frac{n \max_{i=1,\ldots,A} F_i(X_i(0)+1)}{(A-n)\min_{i=1,\ldots,A} F_i(X_i(0))}\right) \xrightarrow{A\to\infty} 1
\end{aligned}$$

□

The same coupling argument also implies, that all agents gain less the more agents there are in the system, i.e.

$$\mathbb{P}\left(X_i^{(A)}(\infty) \geq k\right) > \mathbb{P}\left(X_i^{(A+1)}(\infty) \geq k\right).$$

Another important implication of of Theorem D.1 is a trivial asymptotic independence of agents in large systems.

Corollary D.3. *Assume that (D.1) holds. Then we have for any $A_0 \in \mathbb{N}$ and any Borel-sets $B_1, \ldots, B_{A_0}$ that*

$$\lim_{A\to\infty} \mathbb{P}(X_1^{(A)}(\infty) \in B_1, \ldots, X_{A_0}^{(A)}(\infty) \in B_{A_0}) = \lim_{A\to\infty} \prod_{i=1}^{A_0} \mathbb{P}(X_i^{(A)}(\infty) \in B_i).$$

In order to illustrate Theorem D.1, we executed 100 simulations of a process with $A = 100$ agents and $F_1(k) = \ldots = F_{100}(k) = k^2$. We interrupted the process after 100,000 steps. On average, 30.14 agents won at least one step of the process, with a minimum of 16 and a maximum of 45 agents. In a simulation with $A = 10,000$ agents, interrupted after $10,000,000$ steps, 2010 agents won at least one step. Hence, the convergence in part 2 of Theorem D.1 can be considered as slow.

The following examples show that the assumptions of Theorem D.1 cannot be omitted.

Example D.4.

1. Assume that $\sum_{i=1}^{\infty} F_i(X_i(0)) < \infty$, for example set $F_i(k) = 2^{-i}k^{\beta}$, $\beta > 1$ and $X_i(0) = 1$ for all $i \in \mathbb{N}$, such that (D.1) does not hold. Then the probability that agent 1 wins the first step of the process is

$$\frac{F_1(X_1(0))}{\sum_{i=1}^{A} F_i(X_i(0))} \geq \frac{F_1(X_1(0))}{\sum_{i=1}^{\infty} F_i(X_i(0))} > 0.$$

Thus, $\mathbb{P}\left(X_1^{(A)}(\infty) = X_1(0)\right) \geq \textit{const.}$ does not converge to zero for $A \to \infty$

2. Set $X_i(0) = 1$ for all $i \in \mathbb{N}$. In order to construct an appropriate sequence of feedback functions, take a sequence $(c_i)_i \in (0,1)$, such that $\prod_{i=1}^{\infty} c_i > 0$, e.g. $c_i = e^{-(i^{-2})}$. Moreover, define another sequence $(a_i)_i \in (0,\infty)$ by the recursion formula $a_i \geq \frac{c_i}{1-c_i}\left(\sum_{j=1}^{i-1} a_j + i\right)$, $a_1 = 1$. Then consider the following sequence of feedback functions:

$$F_i(k) = \begin{cases} a_i & \text{if } k = 1 \\ 1 & \text{if } k = 2 \\ k^2 & \text{if } k \geq 3 \end{cases}$$

Note that this sequence fulfills (D.1), but $\limsup_{i\to\infty} F_i(X_i(0)) = \infty$. For any A, the probability that agent i wins the $(A-i+1)$-th step of the process for all $i \geq \frac{A}{2}$ is given by

$$\prod_{i=\lceil A/2\rceil}^{A} \frac{a_i}{\sum_{j=1}^{i} a_j + A - i} \geq \prod_{\lceil A/2\rceil}^{A} \frac{a_i}{\sum_{j=1}^{i} a_j + 2i - i} \geq \prod_{i=\lceil A/2\rceil}^{A} c_i \geq \prod_{i=1}^{\infty} c_i > 0.$$

Hence, $\mathbb{P}\left(\frac{1}{A}\#\{i \in [A]\colon X_i^{(A)}(\infty) = X_i(0)\} < \frac{1}{2}\right)$ is bounded away from zero (uniformly in A).

Nevertheless, condition (D.1) is fulfilled in most generic situations as the following example underlines. Note that (D.1) does not depend on the initial values $X_i(0)$.

Example D.5. 1. Let $F_i(k) = \alpha_i k^{\beta_i}$ for $\beta_i > 0$. Then (D.1) is fulfilled if $\limsup_{i\to\infty} \min\{\alpha_i, \beta_i - 1\} > 0$.

2. Let $F_i(k) = \alpha_i F(k)$ for a function F satisfying (M). In addition, assume that the $\alpha_i > 0$ are a realisation of an independent and identically distributed sequence of random variables. Then (D.1) is fulfilled, too.

D.2 Large systems with linear or sub-linear feedback

In this subsection, we take a look at large systems where no agent satisfies the monopoly condition (M), such that all agents win infinitely many steps. We keep the notation from the previous Subsection D.1 and denote by $\chi_i^{(A)}(n) = X_i^{(A)}(n)/\sum_{j\in[A]} X_j^{(A)}(n)$ the process of shares corresponding to $X_i^{(A)}(n)$. For the classical Pólya urn with linear feedback, it is well known that the long-time market shares

$$\chi_i^{(A)}(\infty) := \lim_{n\to\infty} \chi_i^{(A)}(n) \in [0,1]$$

exist and that $\left(\chi_1^{(A)}(\infty), \dots, \chi_A^{(A)}(\infty)\right)$ has a Dirichlet distribution (Theorem 2.4). The following Theorem states that $X_i^{(A)}(n)$ is approximately exponentially distributed with mean $\frac{n}{A} + 1$, when A and n is large and when all agents start with the same wealth at time zero. Moreover, any fixed number of agents is asymptotically independent for $A \to \infty$. Thus, the empirical common wealth distribution of all agents is asymptotically exponential and the share of the richest agent is approximately given by $(\log A)/A$.

Theorem D.6. *Let $F_1(k) = F_2(k) = \ldots = k$ and $X_1(0) = X_2(0) = \ldots = 1$. Moreover, fix $A_0 \in \mathbb{N}$ and $x_1, \ldots, x_{A_0} > 0$. Then the following holds:*

$$\mathbb{P}\left(A \cdot \chi_1^{(A)}(\infty) \leq x_1, \ldots, A \cdot \chi_{A_0}^{(A)}(\infty) \leq x_{A_0}\right) \xrightarrow{A\to\infty} \prod_{i=1}^{A_0} (1 - e^{-x_i}) \tag{D.3}$$

$$\frac{A}{\log(A)} \max_{i=1,\ldots,A} \chi_i^{(A)}(\infty) \xrightarrow{A\to\infty} 1 \quad \textit{in distribution} \tag{D.4}$$

$$\sup_{x\geq 0} \left| \frac{1}{A} \sum_{i=1}^{A} \mathbb{1}_{\{A\chi_i^{(A)}(\infty) > x\}} - e^{-x} \right| \xrightarrow{A\to\infty} 0 \quad \textit{in distribution} \tag{D.5}$$

Proof. For (D.3), suppose that A is large enough, such that $\sum_{i=1}^{A_0} \frac{x_i}{A} < 1$. It is well known, that $\left(\chi_1^{(A)}(\infty), \ldots, \chi_A^{(A)}(\infty)\right)$ is uniformly distributed on the simplex Δ_{A-1} for all A. Then:

$$\begin{aligned}
&\mathbb{P}\left(A \cdot \chi_1^{(A)}(\infty) \leq x_1, \ldots, A \cdot \chi_{A_0}^{(A)}(\infty) \leq x_{A_0}\right) = \mathbb{P}\left(\chi_1^{(A)}(\infty) \leq \frac{x_1}{A}, \ldots, \chi_{A_0}^{(A)}(\infty) \leq \frac{x_{A_0}}{A}\right) = \\
&= \int_0^{x_1/A} \cdots \int_0^{x_{A_0}/A} \int_0^{1-y_1-\ldots-y_{A_0}} \cdots \int_0^{1-y_1-\ldots-y_{A-1}} (A-1)!\, dy_A \ldots dy_1 \\
&= (A-1)! \int_0^{x_1/A} \cdots \int_0^{x_{A_0}/A} \int_0^{1-y_1-\ldots-y_{A_0}} \cdots \int_0^{1-y_1-\ldots-y_{A-2}} (1 - y_1 - \ldots - y_{A-1})\, dy_{A-1} \ldots dy_1 \\
&= \frac{(A-1)!}{2} \int_0^{x_1/A} \cdots \int_0^{x_{A_0}/A} \int_0^{1-y_1-\ldots-y_{A_0}} \cdots \int_0^{1-y_1-\ldots-y_{A-3}} (1 - y_1 - \ldots - y_{A-2})^2\, dy_{A-2} \ldots dy_1 \\
&= \frac{(A-1)!}{(A-A_0)!} \int_0^{x_1/A} \cdots \int_0^{x_{A_0}/A} (1 - y_1 - \ldots - y_{A_0})^{A-A_0}\, dy_{A_0} \ldots dy_1 \\
&= \frac{(A-1)!}{(A-A_0+1)!} \int_0^{x_1/A} \cdots \int_0^{x_{A_0-1}/A} (1 - y_1 - \ldots - y_{A_0-1})^{A-A_0+1} \\
&\qquad - (1 - y_1 - \ldots - y_{A_0-1} - \frac{x_{A_0}}{A})^{A-A_0+1}\, dy_{A_0-1} \ldots dy_1 \\
&= \sum_{\{i_1,\ldots,i_k\}\subset[A_0]} (-1)^k \left(1 - \frac{x_{i_1}}{A} - \ldots - \frac{x_{i_k}}{A}\right)^A = \\
&\xrightarrow{A\to\infty} \sum_{\{i_1,\ldots,i_k\}\subset[A_0]} (-1)^k \exp\left(-\sum_{l=1}^{k} x_{i_l}\right) = \sum_{\{i_1,\ldots,i_k\}\subset[A_0]} (-1)^k \prod_{l=1}^{k} e^{-x_{i_l}} = \prod_{i=1}^{A_0} \left(1 - e^{-x_i}\right)
\end{aligned}$$

Next, (D.4) follows from the fact that $\chi^{(A)}(\infty) \in \Delta_{A-1}$ has a Dirichlet distribution with parameter $X(0)$, which allows for the following representation from [57, Theorem 1.1]: Let $Z_1, \ldots, Z_A$ be independent, exponentially distributed random variables with expectation 1.

Then

$$\left(\frac{Z_i}{Z_1 + \ldots + Z_A}\right)_{i\in[A]}$$

has also a Dirichlet distribution with parameter $(1,\ldots,1)$, such that (D.4) is easy to check.

Finally, (D.5) follows from [57, Theorem 2.17], which states that in the exponential embedding $e^{-t}\,\Xi_i(t)$ converges for $t\to\infty$ almost surely to an exponentially distributed random value for all $i\in[A]$. Hence, the left hand side of (D.5) is the empirical distribution function of an independent sample of A exponential distributions. Then the claim is a consequence of the Glivenko–Cantelli theorem. □

For $X(0)=(1,\ldots,1)$ and large A, the shares $\chi^{(A)}(n)$ do barely change in each step, such that no agent wins a significant share on the long run. This can be fixed by replacing the state space of the process $X^{(A)}(n)$ by $(0,\infty)^A$, while keeping all definitions of Section 2.1 mutatis mutandis. If $X(0)=\left(\frac{1}{A},\ldots,\frac{1}{A}\right)$, then $\chi^{(A)}(\infty)$ still has a Dirichlet distribution with parameter $X(0)$. According to [57, Theorem 2.1], $\chi^{(A)}(\infty)$ converges for $A\to\infty$ towards a Poisson-Dirichlet distribution (up to sorting).

Finally, we take a quick look at the sub-linear case. As derived in Section 3.2, we know that the long-time market shares $\chi_i^{(A)}(\infty)$ are deterministic and independent of the initial values. In order to bring some randomness to our model, we assume that all agents have the same feedback function up to some random fitness parameter α_i. When A and n is large, then the distribution of $X_i^{(A)}(n)$ is approximately determined by the distribution of the α_i up to scaling and a distortion depending on F. Moreover, the agents are asymptotically independent. The following proposition formalizes this idea.

Proposition D.7. *Let $F_i(k)=\alpha_i F(k)$ for all $i,k\in\mathbb{N}$, where F fulfills*

$$\frac{F_i(k)}{k}\sum_{l=1}^{k}\frac{1}{F_i(l)} \xrightarrow{k\to\infty} c\in(0,\infty)\,.$$

In addition, suppose that $(\alpha_i)_{i\in\mathbb{N}}$ is a sequence of independent, positive and identically distributed random variables with $\mathbb{E}\alpha_i^c\in(0,\infty]$. Fix $A_0\in\mathbb{N}$. Then:

$$\left(A\cdot\chi_1^{(A)}(\infty),\ldots,A\cdot\chi_{A_0}^{(A)}(\infty)\right) \xrightarrow{A\to\infty} \left(\frac{\alpha_1^c}{\mathbb{E}\alpha_1^c},\ldots,\frac{\alpha_{A_0}^c}{\mathbb{E}\alpha_1^c}\right) \quad \textit{in distribution.}$$

Proof. From Corollary 3.29, we know that $\chi_i^{(A)}(\infty)=\left(1+\alpha_i^{-c}\sum_{j\neq i}\alpha_j^c\right)^{-1}$ almost surely. Hence,

$$A\cdot\chi_i^{(A)}(\infty)=\left(\frac{1}{A}+\alpha_i^{-c}\frac{1}{A}\sum_{j\neq i}\alpha_j^c\right)^{-1} \xrightarrow{A\to\infty} \frac{\alpha_i^c}{\mathbb{E}\alpha_1^c} \quad \text{almost surely}$$

due to the law of large numbers. □

In particular, if the α_i have a heavy-tailed distribution, such that $\mathbb{E}\alpha_i^c = \infty$, then $A \cdot \chi_1^{(A)}(\infty) \xrightarrow{A\to\infty} 0$. A possible interpretation of that is that the wealth of a typical agent grows slower than the whole economy. Vice versa, most of the wealth is hold by the most attractive agents.

Now assume that α_i has finite moments and that $F(k) = k^\beta$, $\beta < 1$, such that $c = \frac{1}{1-\beta}$. Let A and n be large. If $\beta = 0$, then the distribution of $X_i^{(A)}(n)$ coincides with the distribution of α_i up to scaling. This means that any distribution of wealth can be obtained only by differences in the attractiveness of the agents and without increasing returns. When β gets closer to one, then the distribution of $X_i^{(A)}(n)$ gets more unequal. This is consistent with the idea of weak monopoly for the linear case with different attractiveness parameters (see Proposition 3.30).

D.3 Dependence of $c(A, a)$ on A and a

In Subsection 5.2.2, we introduced the constant $c(A, a)$ as a measure for the tail dependence of the wealth of a super-linear losers in a system of $A > a$ agents in total. We have already discussed that $c(A, a) > 1$ in a symmetric situation, corresponding to a positive correlation of losers. This section examines the dependence of $c(A, a)$ on A and a, starting with A. We use the notation introduced in Chapter 5, in particular Section 5.2.

$c(A, a)$ can by computed numerically via formula (5.15), since Lemma 5.3 provides explicit expressions for g_i and G_i and, consequently, also for the density of S. Hence, the expectations in (5.15) can be well approximated by Riemann sums. In the symmetric case, the density of S is simply given by $(A-a)G(s)^{A-a-1}g(s)$, $s \geq 0$. The following table shows approximations of $c(A, 2)$ for different feedback functions and system sizes in the symmetric case $F_1 = \ldots = F_A = F$ and $X(0) = (1, \ldots, 1)$. The formula from Lemma 5.3 was truncated after 100 summands. For the Riemann approximation of the expectation integral, we took a bandwidth of 0.0001 and computed the integral on the interval $[0, 50]$ (c.f. Figure 5.2)

c(A, 2)	$A = 3$	$A = 10$	$A = 100$	$A = 10^3$	$A = 10^6$
$F(k) = k^2$	1.121	1.227	1.427	1.565	1.754
$F(k) = e^{k-1}$	1.130	1.218	1.374	1.480	1.634
$F(k) = k(\log(e-1+k))^2$	1.141	1.254	1.460	1.597	1.779

The table confirms our finding $c(A, a) > 1$ from Proposition 5.12. Moreover, we observe that $c(A, 2)$ does not strongly depend on the feedback function, but tends to be larger for slowly increasing F. Remarkably, $c(A, 2)$ seems to be slightly increasing in A. Thus, the tail dependence of agents seems to be stronger in larger systems, as opposed to the classical Pólya urn with linear feedback, where one can show an asymptotic independence of agents for $A \to \infty$ (see Appendix D.1). Heuristically, this phenomenon may be explained as follows:

First, we observe the implication

$$X_2(\infty) \geq x_2 \quad \Longrightarrow \quad S > T_2(x_2) := \sum_{k=X_2(0)}^{x_2-1} \tau_2(k) \tag{D.6}$$

and that the first passage time $T_2(x_2)$ is independent of A and S. Since $S = S(A)$ is decreasing in A, the event $S > T_2(x_2)$ is less likely for larger A. Hence, conditioning the distribution of S on the increasingly unlikely event $X_2(\infty) \geq x_2$ has a stronger effect for larger A. So S is more likely to be untypically large and hence $X_1(\infty)$ is slightly more likely to be large as well. The following Proposition and Conjecture D.9 underline this idea.

Proposition D.8. *Assume $X_1(0) = X_2(0) = \ldots$ and that $F_1 = F_2 = \ldots$ satisfy (M). Define $S = S(A) := \min_{i=3,\ldots,A} T_i$. Then we have for all $s > 0$ and $x_2 > X_2(0)$*

$$\frac{\mathbb{P}_{A,2}(S(A) > s \,|\, X_2(\infty) \geq x_2)}{\mathbb{P}_{A,2}(S(A) > s)} \to \infty \quad \textit{for } A \to \infty\,,$$

but still

$$\mathbb{P}_{A,2}(S(A) > s \,|\, X_2(\infty) \geq x_2) \to 0 \quad \textit{for } A \to \infty\,.$$

Proof. First, we observe for any $0 < s < t$ that

$$\frac{\mathbb{P}(S > s)}{\mathbb{P}(S > t)} = \left(\frac{\mathbb{P}(T_1 > s)}{\mathbb{P}(T_1 > t)}\right)^{A-2} \to \infty \quad \text{and} \quad \mathbb{P}(s < S < t) \sim \mathbb{P}(S > s) \quad \text{for } A \to \infty\,.$$

Hence, we get for any positive random variable T, which is independent of S, that

$$\mathbb{P}(s < S < T) = \int_s^\infty \mathbb{P}(s < S < t)\, dP_T(t) \sim \mathbb{P}(S > s)\, \mathbb{P}(T > s) \quad \text{for } A \to \infty\,, \tag{D.7}$$

where P_T denotes the law of T. As a consequence, we get for all $s \geq 0$

$$P_{A,2}(S > s) = \frac{\mathbb{P}(S > s,\, S < T_1,\, S < T_2)}{\mathbb{P}(S < T_1,\, S < T_2)} \sim \frac{\mathbb{P}(s < T_1,\, s < T_2)}{\mathbb{P}(S < T_1,\, S < T_2)} \mathbb{P}(S > s) \quad \text{for } A \to \infty\,. \tag{D.8}$$

Let $P_{(T_2, T_2(x_2))}$ be the joint law of $(T_2, T_2(x_2))$ with $T_2(x_2)$ as in (D.6). Now, we are prepared for the final calculation, using the dominated convergence theorem we get for all $s \geq 0$ and $x_2 > X_2(0)$

$$
\begin{aligned}
&\frac{\mathbb{P}_{A,2}(S > s \,|\, X_2(\infty) \geq x_2)}{\mathbb{P}_{A,2}(S > s)} = \frac{\mathbb{P}(S > s,\, S > T_2(x_2),\, S < T_1,\, S < T_2)}{\mathbb{P}_{A,2}(S > s)\,\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)} \\
&= \int_0^\infty \int_0^t \frac{\mathbb{P}(S > s,\, S > t(x_2),\, S < T_1,\, S < t) dP_{(T_2, T_2(x_2))}(t, t(x_2))}{\mathbb{P}_{A,2}(S > s)\,\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)} \\
&\geq \int_0^\infty \int_0^t \frac{\mathbb{P}(S > s,\, S < T_1,\, S < t)}{\mathbb{P}_{A,2}(S > s)\,\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)} \cdot \mathbb{1}_{\{t(x_2) < s\}} dP_{(T_2, T_2(x_2))}(t, t(x_2)) \\
&\sim \int_0^\infty \int_0^s \frac{\mathbb{P}(S < T_1,\, S < T_2)\,\mathbb{P}(\min\{T_1,\, t\} > s)}{\mathbb{P}(s < T_1, s < T_2)\,\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)} dP_{(T_2, T_2(x_2))}(t, t(x_2)) \\
&\asymp \frac{\mathbb{P}(S < T_1,\, S < T_2)}{\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)} \to \infty \quad \text{for } A \to \infty \,,
\end{aligned}
$$

using the aymptotics (D.7), (D.8) and

$$
\int_0^\infty \int_0^s \mathbb{P}(\min\{T_1,\, t\} > s) dP_{(T_2, T_2(x_2))}(t, t(x_2)) \asymp \mathbb{P}(s < T_1, s < T_2)\,,
$$

with $\mathbb{P}\big(T_2(x_2) < s\big) \in (0, 1)$. The second part of the claim follows as follows:

$$
\begin{aligned}
&P_{A,2}(S > s \,|\, X_2(\infty) \geq x_2)^{-1} = \frac{\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2)}{\mathbb{P}(S > T_2(x_2),\, S < T_1,\, S < T_2, S > s)} \\
&\geq \int_0^\infty \int_0^t \frac{\mathbb{P}(S > t(x_2),\, S < T_1,\, S < t) \mathbb{1}_{\{t(x_s) < s\}}\, dP_{(T_2, T_2(x_2))}(t, t(x_2))}{\mathbb{P}(S > s)} \\
&\sim \int_0^\infty \int_0^s \frac{\mathbb{P}(S > t(x_2))}{\mathbb{P}(S > s)}\, \mathbb{P}(t(x_2) < \min\{T_1,\, t\})\, dP_{(T_2, T_2(x_2))}(t, t(x_2)) \xrightarrow{A \to \infty} \infty \,.
\end{aligned}
$$

□

Hence, S becomes stochastically larger by conditioning on $X_2(\infty) \geq x_2$ and this change is stronger for larger A. Since S covers the dependence between both agents, $X_1(\infty) = \Xi_1(S)$ also becomes stochastically larger by conditioning on $X_2(\infty) \geq x_2$ and the effect is increasing with A, causing an increasing dependence in the sense measured by $c(A, a)$. Be aware that Proposition D.8 holds for fixed (i.e. non-diverging) x_2. Nevertheless, the dependence of losers is rather weak for all A as the tail decay of $X_i(\infty)$ is only affected by other agents up to a constant prefactor, which is however not negligible even in the limit $A \to \infty$. In general, the $X_i(\infty)$ are asymptotically independent for $A \to \infty$ in a trivial way (see Appendix D.1). This is consistent with the second part of Proposition D.8, which implies for large A that $X_1(\infty) = X_1(0)$ is likely even conditioned on $X_2(\infty) \geq x_2$. This is not a contradiction as the constant $c(A, a)$ only measures the tail dependence of agents who won many steps.

The slightly positive correlation of losers can also be numerically underlined as follows. For $F(k) = k^2$, we executed 100,000 simulations each with $A = 3$ and $A = 30$. The simulations were stopped when one agent exceeded a share of 0.99. The Pearson correlation coefficient of

$\log(X_1(\infty))$ and $\log(X_2(\infty))$ conditioned on $sMon_1^c \cap sMon_2^c$ amounts to $+0.020$ for $A = 3$ and $+0.025$ for $A = 30$.

Next, we analyse the dependence of $c(A,a)$ on a. For that, let us rephrase Theorem 5.11 in the fully symmetric case as

$$\mathbb{P}_{A,a}(X_1(\infty) > x_1 \,|\, X_2(\infty) > x_2, \ldots, X_a(\infty) > x_a) \sim \frac{c(A,a)}{c(A,a-1)} \mathbb{P}_{A,a}(X_1(\infty) > x_1)$$

for $x_1, \ldots, x_a \to \infty$. Note that $c(A,1) = 1$ by definition. Obviously, the information that several losers won in many steps is a stronger hint for a late explosion of the winner than having this information for only a few agents. Hence, we conjure that $c(A,a)/(A,a-1)$ is increasing in a and, equivalently, $c(A,a)$ increases super-exponentially in a. The following table shows the ratio $c(A,a)/(A,a-1)$ for $F(k) = k^2$ and different A, a.

$\frac{c(A,a)}{c(A,a-1)}$	$A = 10^3$	$A = 10^6$	$A = 10^9$
$a = 3$	2.04	2.45	2.61
$a = 10$	4.37	6.65	7.63
$a = 20$	6.45	11.78	14.14
$a = 30$	8.25	16.42	20.27

These computations combined with the heuristic below motivate the following conjecture for large systems.

Conjecture D.9. *For $F_1 = \ldots = F_A$, $X(0) = (1, \ldots, 1)$ and any $a \geq 2$, we have*

$$\lim_{A\to\infty} c(A,a) = a! \quad \textit{so that} \quad \lim_{A\to\infty} \frac{c(A,a)}{c(A,a-1)} = a \ .$$

The precise value $a!$ is justified by the following heuristic argument. First, we observe that $c(A,a) \sim \frac{\mathbb{E}[g(S)^a]}{\mathbb{E}[g(S)]^a}$ for $A \to \infty$ since $S \xrightarrow{A\to\infty} 0$ almost surely. Now, we approximate for simplicity

$$\mathbb{P}(T_i \leq s) \approx cs^k \quad \text{for } s \to 0 \tag{D.9}$$

for some $c > 0$ and k large (see Lemma 5.2). By standard extreme value theory, we know that $\sqrt[k]{A}S$ converges for $A \to \infty$ towards a Weibull distributed random variable W with density $w(s) = cks^{k-1}e^{-cs^k}$. Then we get for $A \to \infty$:

$$\frac{\mathbb{E}[g(S)^a]}{\mathbb{E}[g(S)]^a} \sim \frac{\mathbb{E}[g(W/\sqrt[k]{A})^a]}{\mathbb{E}[g(W/\sqrt[k]{A})]^a} \sim \frac{\mathbb{E}\left[(cW^k/A)^a\right]}{\mathbb{E}[cW^k/A]^a} = \frac{\mathbb{E}[W^{ak}]}{\mathbb{E}[W^k]^a} = \frac{c^{-a}\Gamma(1+a)}{c^{-a}\Gamma(2)^a} = a!$$

Since the limit does neither depend on k nor on c, it stands to reason that the approximation (D.9) does not change the result. If conjecture D.9 holds, then $c(A,a)/c(A,a-1) \to a > 1$, underlining that the positive tail dependence of losers does not vanish in the limit $A \to \infty$.

Appendix E

The Time of Monopoly

As already explained in Section 5.2.3, Theorem 5.10 and Theorem 5.16 are useful to determine the tail distribution of the time of monopoly as defined in (5.18). In the symmetric, super-linear two-agent case, this quantity has already been studied in [132, 44], but with a slightly different definition (which differs from ours at most up to a multiplicative constant). They found that the tail of N_{mon} is heavier than the tail of $\min\{X_1(\infty), X_2(\infty)\}$, which implies that it is not untypical for losers to win a few late steps, when the advantage of the winner is already big. In this chapter, we extend their results to the (mostly homogeneous) asymmetric case, where we also allow for sub-linear feedback.

E.1 Time of monopoly for super-linear agents

We directly formulate our main result for a system with two super-linear agents, who can both be the monopolist with positive probability. We will return to $A > 2$ below.

Corollary E.1. *Let $A = 2$ and $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_i > 1$ and $\alpha_i > 0$. Define $\beta := \frac{\beta_1 - 1}{\beta_2}$ and assume that (E.1) (see below) holds.*

1. *If $\beta_1 \le \beta_2 + 1$, then we have*

$$\mathbb{P}(N_{mon} = n \,|\, sMon_1) \asymp n^{\beta - \beta_1} \quad \text{for } n \to \infty\,.$$

2. *If $\beta_1 \ge \beta_2 + 1$, then we have*

$$\mathbb{P}(N_{mon} = n \,|\, sMon_1) \asymp n^{-\beta_2} \quad \text{for } n \to \infty\,.$$

In the proof, k will denote the number of steps won by agent 2, when the monopoly of agent 1 set in at time n. We will have to distinguish small, moderate and large k for given n. For that, we define

$$k_1(n) := \begin{cases} n^{\beta} & \text{for } \beta_1 < \beta_2 + 1 \\ \frac{n}{2} & \text{for } \beta_1 = \beta_2 + 1 \\ n - n^{\beta_2/(\beta_1 - 1)} & \text{for } \beta_1 > \beta_2 + 1 \end{cases}$$

and

$$k_2(n) := \begin{cases} const.n^{(\beta_1-1)/(\beta_2-1)} & \text{for } \beta_1 < \beta_2 \\ const.\frac{n}{2} & \text{for } \beta_2 \leq \beta_1 < \beta_2 + 1 \\ n - const.n^{\beta_2/\beta_1} & \text{for } \beta_1 \geq \beta_2 + 1 \end{cases}$$

with *const.* > 0, such that (E.2) holds. Note that $k_2(n) \geq k_1(n)$ for all n. In order to apply the dominated convergence theorem in the proof of Corollary E.1, we need to require the following uniformity of the convergence $X_2(n) \to X_2(\infty)$:

$$\sup_{\substack{n\in\mathbb{N} \\ k_1(n)\leq k\leq k_2(n)}} \frac{\mathbb{P}(X_2(n) = X_2(0) + k,\, sMon_1)}{\mathbb{P}(X_2(\infty) = X_2(0) + k)} < \infty \tag{E.1}$$

This assumptions seems plausible since the convergence $X_2(n) \nearrow X_2(\infty)$ is monotone. Nevertheless, a formal proof is still pending. This specific choice of $k_1(n)$ and $k_2(n)$ ensures that

$$\liminf_{n\to\infty} \frac{(n-k_2(n))^{\beta_1}}{k_2(n)^{\beta_2}} > 0, \quad \limsup_{n\to\infty} \frac{k_1(n)^{\beta_2}}{(n-k_1(n))^{\beta_1-1}} < \infty, \quad \liminf_{n\to\infty} \frac{k_2(n)^{\beta_2}}{(n-k_2(n))^{\beta_1-1}} = \infty$$

holds, which we are going to use in the proof of Corollary E.6 extensively. We start the proof with a short lemma.

Lemma E.2. *Let $A = 2$ and $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_i > 1$ and $\alpha_i > 0$. Then:*

$$\liminf_{n\to\infty} \inf_{k_1(n)\leq k\leq k_2(n)} \frac{\mathbb{P}(X_2(n) = X_2(0) + k,\, sMon_1)}{\mathbb{P}(X_2(n) = X_2(0) + k)} > 0$$

Proof. We have to show that

$$\mathbb{P}(sMon_1 \,|\, X_2(n) = X_2(0) + k) \geq \mathbb{P}(sMon_1 \,|\, X_2(n) = X_2(0) + k_2(n))$$

is bounded away from zero. Using the concentration property of explosion times around their expectation from Lemma 3.15, it suffices to show that

$$\liminf_{n\to\infty} \frac{\sum_{l=X_2(0)+k_2(n)}^{\infty} \frac{1}{F_2(l)}}{\sum_{l=X_1(0)+n-k_2(n)}^{\infty} \frac{1}{F_1(l)}} = \lim_{n\to\infty} \frac{\alpha_1(\beta_2-1)}{\alpha_2(\beta_1-1)} \cdot \frac{k_2(n)^{1-\beta_2}}{(n-k_2(n))^{1-\beta_1}} > 1\,, \tag{E.2}$$

which is ensured by the definition of $k_2(n)$ in each case. □

Proof of Corollary E.1. First, we express the desired probability as

$$\begin{aligned}
&\mathbb{P}(N_{mon} = n+1, sMon_1) \qquad \text{(E.3)}\\
&= \sum_{k=1}^{n-1} \mathbb{P}(X_2(n-1) = k + X_2(0)) \cdot \mathbb{P}(N_{mon} = n+1, sMon_1 \,|\, X_2(n-1) = k + X_2(0))\\
&= \sum_{k=1}^{n-1} \mathbb{P}(X_2(n-1) = k + X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k) + F_2(X_2(0)+k)} \cdot R(k,n)\,,
\end{aligned}$$

where

$$R(k,n) := \prod_{l=0}^{\infty} \frac{F_1(X_1(0)+n-k-1+l)}{F_1(X_1(0)+n-k-1+l) + F_2(X_2(0)+k+1)} < 1 \qquad \text{(E.4)}$$

is the probability, that agent 1 wins all steps after step n, when agent 2 won $k+1$ of the first n steps. Due to $\frac{1}{1+x} \geq e^{-x}$ for all $x > -1$, we have the following lower bound for $R(k,n)$:

$$R(k,n) \geq \exp\left(-F_2(X_2(0)+k+1) \sum_{l=X(0)+n-k-1}^{\infty} \frac{1}{F_1(l)} \right) \geq \exp\left(-const.k^{\beta_2}(n-k)^{1-\beta_1} \right)$$

1. Now assume that $\beta_1 - \beta_2 < 1$, such that $\beta < 1$ and $R(k,n)$ is uniformly bounded away from zero as long as $k \leq k_1(n) = n^\beta$. Using Theorem 5.10, we get the following asymptotic for the first $k_1(n)$ summands in (E.3):

$$\begin{aligned}
&\sum_{k=1}^{k_1(n)} \mathbb{P}(X_2(n-1) = k + X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k) + F_2(X_2(0)+k)} \cdot R(k,n)\\
&\asymp \sum_{k=1}^{n^\beta} \frac{\mathbb{P}(X_2(n-1) = k + X_2(0))}{\mathbb{P}(X_2(\infty) = k + X_2(0))} \cdot \frac{1}{F_1(n-k)} \asymp n^{-\beta_1} \sum_{k=1}^{n^\beta} \frac{\mathbb{P}(X_2(n-1) = k + X_2(0))}{\mathbb{P}(X_2(\infty) = k + X_2(0))}\\
&= n^{-\beta_1} \int_0^{n^\beta} \frac{\mathbb{P}(X_2(n-1) = \lceil u \rceil + X_2(0))}{\mathbb{P}(X_2(\infty) = \lceil u \rceil + X_2(0))}\, du = n^{\beta-\beta_1} \int_0^1 \frac{\mathbb{P}(X_2(n-1) = \lceil n^\beta u \rceil + X_2(0))}{\mathbb{P}(X_2(\infty) = \lceil n^\beta u \rceil + X_2(0))}\, du\\
&\sim n^{\beta-\beta_1} \quad \text{for } n \to \infty \qquad \text{(E.5)}
\end{aligned}$$

In the last line, we applied the dominated convergence theorem, which holds since

$$\begin{aligned}
&\frac{\mathbb{P}(X_2(n-1) = \lceil n^\beta u \rceil + X_2(0))}{\mathbb{P}(X_2(\infty) = \lceil n^\beta u \rceil + X_2(0))} \leq \frac{\mathbb{P}(X_2(n-1) = \lceil n^\beta u \rceil + X_2(0))}{\mathbb{P}(X_2(\infty) = X_2(n-1) = \lceil n^\beta u \rceil + X_2(0))}\\
&= \frac{1}{\mathbb{P}(X_2(\infty) = X_2(n-1) \,|\, X_2(n-1) = \lceil n^\beta u \rceil + X_2(0))} = \frac{1}{R(\lceil n^\beta u \rceil - 1, n-1)}\\
&\leq \frac{1}{R(\lceil n^\beta \rceil - 1, n-1)} \qquad \text{(E.6)}
\end{aligned}$$

To complete the proof of 1., it remains to show that the remaining summands in (E.3) are at most of the same order. For that, we first observe that for $k \leq k_2(n)$

$$
\begin{aligned}
R(k,n) &\leq \exp\left(-const.F_2(X_2(0)+k+1)\sum_{l=X(0)+n-k-1}^{\infty}\frac{1}{F_1(l)}\right) \\
&\leq \exp\left(-const.k^{\beta_2}(n-k)^{1-\beta_1}\right)
\end{aligned}
\tag{E.7}
$$

holds due to $\frac{1}{1+x} \leq e^{-cx}$ for $c \leq \frac{\log(1+x)}{x}$ and $x > 0$. Now, we have to split the sum in (E.3) once again. Now, we need Lemma E.2 and (E.1).

$$
\begin{aligned}
&\sum_{k=k_1(n)+1}^{k_2(n)} \mathbb{P}(X_2(n-1)=k+X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k)+F_2(X_2(0)+k)} \cdot R(k,n) \qquad \text{(E.8)} \\
&\prec \sum_{k=n^{\beta}+1}^{k_2(n)} \frac{1}{F_1(n)} \exp\left(-const.k^{\beta_2}(n-k)^{1-\beta_1}\right) \prec n^{-\beta_1} \sum_{k=n^{\beta}+1}^{k_2(n)} \exp\left(-const.k^{\beta_2}n^{1-\beta_1}\right) \\
&\leq n^{-\beta_1} \int_{n^{\beta}}^{k_2(n)} \exp\left(-const.u^{\beta_2}n^{1-\beta_1}\right) du = n^{-\beta_1}n^{\beta} \int_{1}^{k_2(n)n^{-\beta}} \exp\left(-const.u^{\beta_2}\right) du \\
&\asymp n^{-\beta_1}n^{\beta} \quad \text{for } n \to \infty
\end{aligned}
$$

For the remaining summands in (E.3), a rough estimate is sufficient:

$$
\begin{aligned}
&\sum_{k=k_2(n)+1}^{n-1} \mathbb{P}(X_2(n-1)=k+X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k)+F_2(X_2(0)+k)} \cdot R(k,n) \\
&\leq \sum_{k=k_2(n)+1}^{n-1} R(k,n) \leq \sum_{k=k_2(n)+1}^{n-1} R(k_2(n),n) \leq (n-k_2(n)) \exp\left(-const.k_2(n)^{\beta_2}n^{1-\beta_1}\right) \qquad \text{(E.9)}
\end{aligned}
$$

Obviously, the last term decays faster than any polynomial for $\beta_1 - \beta_2 < 1$.

2. Let $\beta_1 - \beta_2 \geq 1$. Since $N_{mon} \geq X_2(\infty) - X_2(0)$ on the event $sMon_1$, an appropriate asymptotic lower bound follows directly from Theorem 5.10. For an upper bound, define $\beta' = \frac{\beta_2}{\beta_1 - 1}$. Widely analogous to part 1, we have to split the sum (E.3) in three parts again. Suppose for a moment $\beta_1 - \beta_2 > 1$, such that $\beta' < 1$. Then:

$$
\begin{aligned}
&\sum_{k=1}^{k_1(n)} \mathbb{P}(X_2(n-1)=k+X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k)+F_2(X_2(0)+k)} \cdot R(k,n) \\
&\prec \sum_{k=1}^{n-n^{\beta'}} \frac{1}{F_1(n-k)} = \sum_{k=n^{\beta'}}^{n-1} \frac{1}{F_1(k)} \asymp \left(n^{\beta'}\right)^{1-\beta_1} = n^{-\beta_2} \quad \text{for } n \to \infty
\end{aligned}
$$

In the first step, we used (E.6), Theorem 5.10 and $R(k,n) < 1$. Similarly to (E.8), the second partial sum is:

$$
\begin{aligned}
&\sum_{k=k_1(n)+1}^{k_2(n)} \mathbb{P}(X_2(n-1)=k+X_2(0)) \cdot \frac{F_2(X_2(0)+k)}{F_1(X_1(0)+n-1-k)+F_2(X_2(0)+k)} \cdot R(k,n) \\
&\prec \sum_{k=n-n^{\beta'}+1}^{n-n^{\beta_2/\beta_1}} \frac{1}{F_1(n-k)} \exp\left(-const.k^{\beta_2}(n-k)^{1-\beta_1}\right) \\
&\prec \int_{n-n^{\beta'}}^{n-n^{\beta_2/\beta_1}} (n-u)^{-\beta_1} \exp\left(-const.u^{\beta_2}(n-u)^{1-\beta_1}\right) du \\
&= \int_{n^{\beta_2/\beta_1}}^{n^{\beta'}} u^{-\beta_1} \exp\left(-const.(n-u)^{\beta_2}u^{1-\beta_1}\right) du \prec \int_0^{n^{\beta'}} u^{-\beta_1} \exp\left(-const.n^{\beta_2}u^{1-\beta_1}\right) du \\
&= \int_0^1 n^{\beta'-\beta_1\beta'} u^{-\beta_1} \exp\left(-const.u^{1-\beta_1}\right) du = n^{-\beta_2} \int_1^\infty u^{\beta_1-2} \exp\left(-const.u^{\beta_1-1}\right) du
\end{aligned}
$$

In the second line, we used assumption (E.1). The remaining summands decay exponentially fast like in (E.9). For the case $\beta_1 - \beta_2 = 1$, the proof is completely analogous with the given choice of k_1 and k_2. □

Without assumption (E.1), Corollary E.1 still holds if $\asymp$ is replaced by $\succ$ due to first partial sum (E.5). By using Lemma E.2 and the rough estimate

$$
\begin{aligned}
&\mathbb{P}(X_2(n-1)=k+X_2(0)) \leq \mathbb{P}(X_2(n-1) \geq k+X_2(0)) \\
&\asymp \mathbb{P}(X_2(n-1) \geq k+X_2(0), sMon_1) \leq \mathbb{P}(X_2(\infty) \geq k+X_2(0), sMon_1)
\end{aligned} \tag{E.10}
$$

in the middle sum (E.8), we obtain the upper bounds

$$\mathbb{P}(N_{mon}=n \,|\, sMon_1) \prec n^{2\beta-\beta_1} \quad \text{for } \beta_1 \leq \beta_2+1\,,$$

and

$$\mathbb{P}(N_{mon}=n \,|\, sMon_1) \prec n^{1-\beta_2} \quad \text{for } \beta_1 \geq \beta_2+1$$

without assumption (E.1). This upper bound is particularly good for large β_2.

Moreover, it is possible to extend Corollary E.1 to $A > 2$ via the exponential embedding.

Corollary E.3. *Let $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_i > 1$ and $\alpha_i > 0$, $i \in [A]$. Assume $\beta_2 \leq \beta_j$ for all $j \geq 2$ and that (E.1) holds for any two agent system with feedback F_1, F_j. Define $\beta := \frac{\beta_1-1}{\beta_2}$.*

1. *If $\beta_1 \leq \beta_2+1$, then we have*

$$\mathbb{P}(N_{mon} > n \,|\, sMon_1) \asymp n^{\beta-\beta_1+1} \quad \textit{for } n \to \infty\,.$$

2. *If $\beta_1 \geq \beta_2+1$, then we have*

$$\mathbb{P}(N_{mon} > n \,|\, sMon_1) \asymp n^{-\beta_2+1} \quad \textit{for } n \to \infty\,.$$

Proof. First, a lower bound follows from Corollary E.1 via canonical coupling. In the exponential embedding, denote by $s_i \geq 0$, $i \geq 2$ the time of the last jump of Ξ_i before the explosion time T_1 of agent 1. Then, N_{mon} can be expressed as

$$N_{mon} = \Xi_1 \left(\max_{i=2,\ldots,A} s_i \right) + X_2(\infty) + \ldots X_A(\infty) + 1$$

on the event $sMon_1$. Hence:

$$\mathbb{P}(N_{mon} > n \,|\, sMon_1) \leq \frac{1}{\mathbb{P}(sMon_1)} \mathbb{P}\left(\sum_{i=2}^{A} (X_i(\infty) + \Xi_1(s_i)) \geq n,\, sMon_1 \right)$$
$$\leq \frac{1}{\mathbb{P}(sMon_1)} \sum_{i=2}^{A} \mathbb{P}\left(X_i(\infty) + \Xi_1(s_i) \geq \frac{n}{A-1},\, T_i > T_1 \right)$$

These probabilities are covered for $n \to \infty$ by Corollary E.1 as they correspond to the probability of $N_{mon} > \frac{n}{A+1}$ in a two agent system. Note that the summand $i = 2$ dominates the others due to the assumption $\beta_2 \leq \beta_j$. □

Hence, in large systems only the loser with the weakest feedback determines the tail of N_{mon}.

Note that the tail weight of the time on monopoly does also depend on the feedback of the winner, as opposed to the wealth of the losers discussed in Theorem 5.10. As a consequence of Corollary E.1, the tail of N_{mon} (without conditioning on a winner) is given by $\mathbb{P}(N_{mon} = n) \asymp n^{\beta - \beta_1}$, where w.l.o.g. $\beta_1 \leq \beta_2$ and $A = 2$. For the symmetric case $\beta_1 = \beta_2$, this is consistent with [132, 44]. According to Corollary E.1, the (unconditioned) tail of N_{mon} is heavier than the one of $\min\{X_1(\infty), X_2(\infty)\}$ in any case. In the conditioned situation, if the feedback of the winner is much stronger than the one of the loser, then N_{mon} is of the same order as the wealth of the loser, i.e. the loser basically only wins steps at the beginning of the process. If the feedback of the loser is at most slightly stronger, then N_{mon} is of higher order than the wealth of the loser and, hence, the loser might win some late steps when the advantage of the winner is already large (see Figure 5.4 for an illustration). Let us formalize this idea.

Corollary E.4. *Let $A = 2$ and $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_i > 1$ and $\alpha_i > 0$. Assume that (E.1) holds. Then:*

1. *If $\beta_1 < \beta_2 + 1$, then*

$$\forall \epsilon > 0\colon \lim_{n \to \infty} \mathbb{P}(\chi_2(N_{mon}) < \epsilon \,|\, N_{mon} = n, sMon_1) = 1\,.$$

2. If $\beta_1 = \beta_2 + 1$*, then*

$$\forall \delta > 0 \; \exists \epsilon > 0 \colon \limsup_{n\to\infty} \mathbb{P}(\chi_2(N_{mon}) < \epsilon \,|\, N_{mon} = n, sMon_1) < \delta$$

and

$$\forall \epsilon > 0 \colon \liminf_{n\to\infty} \mathbb{P}(\chi_2(N_{mon}) > 1 - \epsilon \,|\, N_{mon} = n, sMon_1) > 0 \,.$$

3. If $\beta_1 > \beta_2 + 1$*, then*

$$\forall \epsilon > 0 \colon \lim_{n\to\infty} \mathbb{P}(\chi_2(N_{mon}) > 1 - \epsilon \,|\, N_{mon} = n, sMon_1) = 1 \,.$$

Proof. Recall the proof of Corollary E.6. In particular, we showed that the sum (E.3) is dominated by the summands $k = 1, \ldots, k_2(n)$, where k represents the number of steps won by agent 1, when $N_{mon} = n$ and $sMon_1$. If $\beta_1 < \beta_2 + 1$, then the summands $k = 1, \ldots, \epsilon n$ dominate the whole sum for any $\epsilon > 0$, which implies 1.. In contrast, if $\beta_1 = \beta_2 + 1$, then the summands $k = \epsilon n, \ldots, n$ are of the same order as the whole sum and for any sequence $(l_n)_n$ with $l_n/n \to 0$ the summands $k = l_n, \ldots, n$ dominate the whole sum. Hence, 2. holds. Finally for $\beta_1 > \beta_2 + 1$, the summands $k = 1, \ldots, (1-\epsilon)n$ are of order $n^{1-\beta_1} < n^{-\beta_2}$, such that 3. follows. □

In other words, if the monopoly sets in late, then this typically happens in two different ways depending on the feedback functions: First, if the feedback of the loser is weak compared to the winner, then the loser wins a significant share of steps until the monopoly suddenly set in. For $\beta_1 > \beta_2 + 1$, the loser's share at time N_{mon} is even close to one if the monopoly sets in late. Second, if the feedback of the loser is strong enough and the monopoly sets in late, then the loser can win steps at late time, when the winner was already dominant.

In the situation of 1. in Corollary E.1, the tail of N_{mon} is heavier the weaker the feedback of both agents. This appears intuitive for the winner, but also weak feedback of the loser increases the probability of a late occurrence of monopoly. This surprising fact corresponds to the loser-paradox explained in Section 5.2. The difference between the tail weight of the loser's wealth and of N_{mon} is largest for strong feedback of the loser, since the tail of $X_2(\infty)$ is lighter. These findings are underlined by Figure 5.3 (a). Note that Corollary E.1 is consistent at the transition point $\beta_1 = \beta_2 + 1$. This transition does also occur in the context of total monopoly as discussed in Example 3.9.

For symmetric, exponentially increasing feedback, the tail of N_{mon} decays exponentially as shown in [44]. Let us finally take a quick look at mixed feedback, say $F_1(k) = k^{\beta_1}$ and $F_2(k) = e^{\beta_2 k}$ with $\beta_1 > 1$ and $\beta_2 > 0$. Of course, we still have the trivial lower bounds

$$\mathbb{P}(N_{mon} = n \,|\, sMon_2) \succ n^{-\beta_1} \quad \text{and} \quad \mathbb{P}(N_{mon} = n \,|\, sMon_1) \succ e^{-\beta_2 n}$$

due to Theorem 5.10. Extending 2. of Corollary E.1 to the limit $\beta_2 \to \infty$, it stands to reason that

$$\mathbb{P}(N_{mon} = n \mid sMon_2) \asymp n^{-\beta_1} .$$

The case $sMon_1$ behaves rather surprisingly.

Proposition E.5. *Let $F_1(k) = k^{\beta_1}$ and $F_2(k) = e^{\beta_2 k}$ with $\beta_1 > 1$ and $\beta_2 > 0$. Then:*

$$\mathbb{P}(N_{mon} = n \mid sMon_1) \asymp n^{-\beta_1} \log(n)$$

Proof. Analogously to the proof of Corollary E.6, we consider the sum (E.3) and define $k_1(n) = \log(n^{\beta_1 - 1})/\beta_2$ and $k_2(n) = \log(n^{\beta_1})/\beta_2$. For the first partial sum, we use that $R(k, n)$ (defined mutatis mutandis to (E.4)) is bounded from below.

$$\sum_{k=1}^{k_1(n)} \mathbb{P}(X_2(n-1) = k + X_2(0)) \cdot \frac{F_2(X_1(0)+k)}{F_1(X_2(0)+n-1-k) + F_2(X_1(0)+k)} \cdot R(k,n)$$
$$\asymp \sum_{k=1}^{\log(n^{\beta_1-1})/\beta_2} \frac{1}{(n-k)^{\beta_1}} \sim n^{-\beta_1} \log(n^{\beta_1-1})/\beta_2 \quad \text{for } n \to \infty$$

The dominated convergence theorem is applicable by an analogous argument as (E.6). Repeating the idea of (E.10), we get for the second partial sum:

$$\sum_{k=k_1(n)}^{k_2(n)} \mathbb{P}(X_2(n-1) = k + X_2(0)) \cdot \frac{F_2(X_1(0)+k)}{F_1(X_2(0)+n-1-k) + F_2(X_1(0)+k)} \cdot R(k,n)$$
$$\prec \sum_{\log(n^{\beta_1-1})/\beta_2}^{\log(n^{\beta_1})/\beta_2} \frac{k}{(n-k)^{\beta_1}} e^{-const.e^k (n-k)^{1-\beta_1}} \asymp n^{-\beta_1} \int_{\log(n^{\beta_1-1})/\beta_2}^{\log(n^{\beta_1})/\beta_2} u e^{-const.e^u n^{1-\beta_1}} du$$
$$\prec n^{-\beta_1} \int_1^{\infty} \frac{\log(n^{\beta_1-1}u)}{u} e^{-const.u} du \asymp n^{-\beta_1} \log(n^{\beta_1-1}) \quad \text{for } n \to \infty$$

The remaining summands decay exponentially like in (E.9). □

Basically, this is consistent with the limit $\beta_1 \to \infty$ in part 1. of Corollary E.1, but the additional term $\log(n)$ was not predicted. In particular in the case $sMon_1$, the time of monopoly is heavy-tailed, although the wealth of the loser is light-tailed. Surprisingly, the tail of N_{mon} is heavier on $sMon_1$ than on $sMon_2$. Overall, the (unconditioned) tail of N_{mon} is of order $\log(n)$ heavier than the tail of $\min\{X_1(\infty), X_2(\infty)\}$. Figure 5.3 (b) illustrates this situation. On $sMon_2$, the tails of $X_1(\infty)$ and N_{mon} even seem to be equal, i.e. the loser wins most steps before time N_{mon} and 3. of Corollary E.4 holds analogously.

E.2 Time of monopoly for sub-linear agents

The time of monopoly N_{mon} (see (5.18)) is also well-defined when only one agents satisfies the monopoly condition (M), such that this agent is almost surely the strong monopolist (see

Section 5.3). The following corollary establishes the tail behaviour of N_{mon} in this situation.

Corollary E.6. *Let $A = 2$ and $F_i(k) = \alpha_i k^{\beta_i}$ with $\beta_1 \le 1 < \beta_2$ and $\alpha_i > 0$. If $\beta_1 = 1$, assume additionally $X_2(0) > \beta_2$. Then:*

$$\mathbb{P}(N_{mon} > n) \sim \mathbb{E}F_1(X_1(\infty)) \cdot \sum_{k=n}^{\infty} \frac{1}{F_2(k)} \quad \textit{for } n \to \infty \tag{E.11}$$

Note that $\mathbb{E}F_1(X_1(\infty)) < \infty$ is always satisfied for $\beta_1 < 1$ and is equivalent to $X_2(0) > 1$ for $\beta = 1$ due to Theorem 5.16.

Proof. The probability of $N_{mon} = n + 1$ can be expressed as

$$\begin{aligned}\mathbb{P}(N_{mon} = n+1) = &\sum_{k=0}^{n-1} \mathbb{P}\left(X_1(n-1) - X(0) = k\right)\\ &\cdot \mathbb{P}\left(X(n) - X(n-1) = e^{(1)}, \forall m > n \colon X(m) - X(m-1) = e^{(2)} \,\middle|\, X_1(n-1) - X(0) = k\right)\end{aligned}$$

Since for all $k \ge 0$

$$\lim_{n\to\infty} \mathbb{P}\left(X_1(n-1) - X(0) = k\right) = \mathbb{P}\left(X_1(\infty) - X(0) = k\right)$$

and

$$\begin{aligned}&\mathbb{P}\left(X(n) - X(n-1) = e^{(1)}, \forall m > n \colon X(m) - X(m-1) = e^{(2)} \,\middle|\, X_1(n-1) - X(0) = k\right)\\ &= \frac{F_1(X_1(0)+k)}{F_1(X_1(0)+k) + F_2(X_2(0)+n-k-1)} \cdot \prod_{m>n} \frac{F_2(X_1(0) - 2 - k + m)}{F_1(X_1(0)+k+1) + F_2(X_2(0) - k - 2 + m)}\\ &\sim \frac{F_1(X_1(0)+k)}{F_2(n)} \quad \text{for } n \to \infty\end{aligned} \tag{E.12}$$

holds, we get the following asymptotic lower bound for $n \to \infty$ via Fatou's Lemma:

$$\mathbb{P}(N_{mon} = n+1) \succ \left(\sum_{k=0}^{\infty} \mathbb{P}\left(X_1(\infty) - X(0) = k\right) F_1(X_1(0)+k)\right) \frac{1}{F_2(n)} = \frac{\mathbb{E}F_1(X_1(\infty))}{F_2(n)}$$

Due to non-uniformity of the summands, these considerations are not sufficient for an upper bound, but instead we can split the sum as follows. Take $\epsilon \in (0,1)$, then:

$$\mathbb{P}(N_{mon} = n+1) \leq \mathbb{P}(X_1(\infty) - X_1(0) > \lfloor \epsilon n \rfloor) + \sum_{k=0}^{\lfloor \epsilon n \rfloor} \mathbb{P}\left(X_1(n-1) - X(0) = k\right)$$
$$\cdot \frac{F_1(X_1(0)+k)}{F_1(X_1(0)+k) + F_2(X_2(0)+n-k-1)} \prod_{m>n} \frac{F_2(X_1(0)-2-k+m)}{F_1(X_1(0)+k+1) + F_2(X_2(0)-k-2+m)}$$
$$\leq o(n^{-\beta_2}) + \sum_{k=0}^{\lfloor \epsilon n \rfloor} \mathbb{P}\left(X_1(n-1) - X(0) = k\right) \cdot \frac{F_1(X_1(0)+k)}{F_1(X_1(0)+1) + F_2(X_2(0)+n-\lfloor \epsilon n \rfloor - 1)}$$
$$\leq o(n^{-\beta_2}) + \frac{\mathbb{E}F_1(X_2(n-1))}{F_1(X_1(0)+1) + F_2(X_2(0)+n-\lfloor \epsilon n \rfloor - 1)}$$
$$\sim \frac{1}{(1-\epsilon)^{\beta_2}} \cdot \frac{\mathbb{E}F_1(X_1(\infty))}{F_2(n)} \quad \text{for } n \to \infty$$

In the second inequality, we used Theorem 5.16 and the assumption $X_2(0) > \beta_2$ in the case of $\beta_1 = 1$. $\epsilon \to 0$ finally yields the desired upper bound. □

By an analogous but more lengthy proof, one could generalize Corollary E.6 to $A > 2$. Assume that there are several agents i fulfilling the same assumptions as agent 1 in Corollary E.6, but only one like agent 2. Then (E.11) still holds when $\mathbb{E}F_1(X_1(\infty))$ is replaced by $\sum_i \mathbb{E}F_i(X_i(\infty))$. An analogous extension to exponential F_2 fails at step (E.12). Moreover for almost linear feedback, e.g. $F_1(k) = k\log(k)^{\beta_1}$, $\beta \in (0,1)$, the expectation $\mathbb{E}F_1(X_1(\infty)) = \infty$ is infinite (see Example 5.18), such that an analogous argument is not possible.

As a consequence of Corollary E.6, the time of monopoly N_{mon} has a power law tail, although the tail of $X_1(\infty)$ is lighter than a power-law for $\beta_1 < 1$. Also for $\beta_1 = 1$, one can easily calculate that the tail of N_{mon} is heavier than the one of $X_1(\infty)$. In contrast to $X_1(\infty)$, the tail of N_{mon} does never depend on initial values (up to the constant prefactor). Remarkably in Corollary E.6, the tail decay of N_{mon} does only depend on the feedback of the winner up to a constant prefactor, as opposed to the super-linear case of Appendix E.1. As intuitively expected, the tail decays faster for strong feedback of the winner. All findings of this section are illustrated by Figure 5.5. Let us finally discuss an example to understand the transition between Corollary E.6 and Corollary E.1.

Example E.7. Let $F_i(k) = k^{\beta_i}$ with $\beta_1 > 1$ and varying $\beta_2 \in \mathbb{R}$. Then $\mathbb{P}(N_{mon} > x \,|\, sMon_1) \asymp x^{-a}$ for some exponent $a = a(\beta_2) > 0$. This exponent is constant for $\beta_2 \leq 1$ and jumps downwards (even to zero if $\beta_1 \geq 2$) at $\beta_2 = 1$ in a left-continuous manner. For $\beta_2 > 1$, $a(\beta_2)$ is strictly increasing and with $\lim_{\beta_2 \to \infty} a(\beta_2) = a(0) = 1 - \beta_1$. The tail of N_{mon} coincides with the tail of $X_2(\infty)$ only for $1 < \beta_2 \leq \beta_1 - 1$ (up to constant prefactors). The situation is illustrated by Figure 5.4.