On the Duration and Intensity of Cumulative Advantage Competitions

Bo Jiang,[1] Liyuan Sun,[2] Daniel R. Figueiredo,[3] Bruno Ribeiro,[4] and Don Towsley[1]

[1] *University of Massachusetts, Amherst, MA 01003, USA*
[2] *Tsinghua University, Beijing 100084, China*
[3] *Federal University of Rio de Janeiro, Rio de Janeiro, 21941-972, Brazil*
[4] *Purdue University, West Lafayette, IN 47907, USA*

Network growth can be framed as a competition for edges among nodes in the network. As with various other social and physical systems, skill (fitness) and luck (random chance) act as fundamental forces driving competition dynamics. In the context of networks, cumulative advantage (CA) – the rich-get-richer effect – is seen as a driving principle governing the edge accumulation process. However, competitions coupled with CA exhibit non-trivial behavior and little is formally known about duration and intensity of CA competitions. By isolating two nodes in an ideal CA competition, we provide a mathematical understanding of how CA exacerbates the role of luck in detriment of skill. We show, for instance, that when nodes start with few edges, an early stroke of luck can place the less skilled in the lead for an extremely long period of time, a phenomenon we call "struggle of the fittest". We prove that duration of a simple skill and luck competition model exhibit power-law tails when CA is present, regardless of skill difference, which is in sharp contrast to the exponential tails when fitness is distinct but CA is absent. We also prove that competition intensity is always upper bounded by an exponential tail, irrespective of CA and skills. Thus, CA competitions can be extremely long (infinite mean, depending on fitness ratio) but almost never very intense. The theoretical results are corroborated by extensive numerical simulations. Our findings have important implications to competitions not only among nodes in networks but also in contexts that leverage socio-physical models embodying CA competitions.

I. INTRODUCTION

Growth is a fundamental aspect inherent to most networks that has been widely investigated both empirically through the analysis of data from various contexts and theoretically through idealized models. An important driving force behind network growth and in particular the evolution of node degrees is *cumulative advantage* (CA), where accumulated edges (i.e., current node degree) promote gathering even more edges. A second widely-accepted driving force in this context is *fitness*, which captures the inherent ability of nodes to attract edges. Thus, dynamics of network growth is governed by skill (fitness) and luck (random but biased edge attachment).

Recent work has framed the problem of network growth as a competition among nodes [1–3]. In essence, nodes in a network *compete* with one another to accumulate edges, increasing (or descreasing) their degrees over time. As expected, such competitions are also driven by skill and luck and have been studied empirically and theoretically for different networks, an example of which is the evolution and predictability of success in citation networks [2]. Outside the domain of networks, the study of skill and luck competitions have a long history in social and physical sciences [4, 5].

However, the intricacies of skill and luck competitions are far from trivial, even in a simple CA model with just two competitors. To illustrate, consider a network with two hub nodes that compete for connectivity. Each time a new node joins the network, it connects to one of the hubs randomly with bias depending on the hubs' fitnesses (model details given in Section III). In the presence of CA, the bias also depends on the hubs' current degrees. Figure 1a and Figure 1b illustrate the difference of the hubs' degrees over time for two sample paths in the absence and presence of CA, respectively. The paths with the same color and label in both plots are generated using the same pseudorandom sequence. The competition is tied every time the degree difference is zero and we define the competition duration as the time until the final tie occurs. Note that when generated by the same pseudorandom sequence, a competition with CA can last either much shorter (see the red paths in Figure 1) or *significantly longer* (see the blue paths in Figure 1) than without CA, pointing to a large variance in competition duration. Moreover, with CA the less fit hub enjoys a sizable degree leadership for a long time. These observations also apply to two specific nodes in more general network growth models, provided that we interpret time as the total degrees of the two nodes of interest. However, are these sample paths anomalies or the norm? Can we be more precise about these observations?

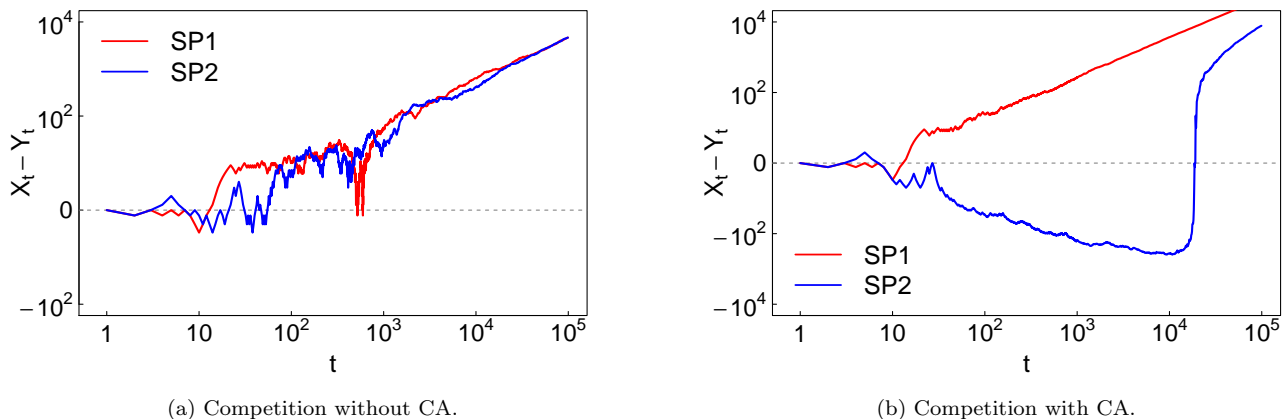(a) Competition without CA. (b) Competition with CA.

FIG. 1.   Time evolution of degree difference between two nodes in competitions without and with cumulative advantage (CA). Each plot shows two independently simulated sample paths. The sample paths with the same color in both plots use the same sequence of pseudorandom numbers. Competition starts tied and node X is 10% fitter than node Y. See model details in Section III.

In this work, we aim to develop a fundamental understanding of the effects of CA in competitions. We approach this problem by considering classical, simple and well-studied theoretical models for competitions based on skill and luck that are either coupled with or free of cumulative advantage. These models may not be general enough as statistical models that fit real-world data for competitions in growing networks, as such models must capture intricate features of the domain, such as skill distribution or amplitude of cumulative advantage (e.g., linear or sub-linear), as well as their time dependency. However, they still provide invaluable insights into how CA impacts competitions. More specifically, we focus on competitions between two agents (nodes) and study two fundamental aspects of competitions: *duration* – the time required for the most skilled to overtake its competitor and forever enjoy undisputed leadership; *intensity* – the number of times competitors tie for the leadership. In this direction, we make the following main

contributions.

- In the case where the two competitors have equal fitness, we obtain the asymptotic tail distributions for both duration and intensity of CA competitions. We demonstrate that they are power laws with respective tail exponents $-1/2$ and -1, which are independent of the initial wealth of the competitors.

- In the case where the two competitors have unequal fitness, we derive asymptotic lower and upper bounds for the tail distribution of duration of CA competitions, and an upper bound for the tail distribution of their intensity. These bounds show that duration is heavy tailed while intensity is exponential tailed in the presence of CA. In particular, duration is heavier tailed while intensity is lighter tailed than corresponding RW competitions.

- We observe that a slight difference in fitness of the two results in an extremely heavy tail for duration of CA competitions. Thus, an individual that is only slightly more skilled than his competitor might have to hang on to the competition for an extremely long period of time before taking the ultimate lead, a phenomenon we call the "struggle of the fittest".

Despite 90 years since the basic CA model was first proposed [6], known as Pólya's urn model, our work is, to the best of our knowledge, the first to characterize the duration and intensity distributions of CA competitions with skill. We believe our findings have profound implications to our understanding of competitions, beyond its importance to the evolution of degree of nodes in growing networks.

The rest of the paper is organized as follows. Section II briefly discusses the related work. Section III introduces the CA competition model. Section IV presents the theoretical results, illustrated and supplemented by simulations. Section V concludes the paper.

II.   RELATED WORK

Resource accumulation is a ubiquitous phenomenon that naturally arises in a variety of social and complex systems. The problem is usually framed as a competition among agents for resources that are abundant, and has been studied in different contexts across various disciplines ranging from protein binding within a cell [7, 8] to views of online social media [9, 10] and citations among scholarly papers [2, 11]. In the context of networks, network growth and in particular node degree evolution has been recently framed as competition among nodes, where different aspects of competitions have been studied empirically and theoretically [1–3, 12, 13].

Models for resource accumulation competitions generally incorporate skill (fitness), luck (randomness) and externalities. Cumulative advantage is one type of externality that is considered a general mechanism for inequality [5]. It appears in the literature under many variants such as Price's cumulative advantage model [11], preferential attachment [14–16], "the rich get richer", Matthew effect [5, 17, 18], and path-dependent increasing returns [4]. The Pólya's urn model [6, 19] is widely used to capture these effects. Most previous work on Pólya's urn model and its generalizations focuses on the share of resources gathered by each agent, also known as the agent's *market share*, proving convergence and limiting results of the market share distribution [19–21]. More recent studies consider Pólya's urn models with non-linear bias [22], the effects of initial conditions of urns [23, 24], as well as the time for the first tie [25] and probability of a tie ever occuring [26].

However, two fundamental metrics associated with competitions, *duration* - how long it takes for the undisputed winner to emerge, and *intensity* - how many times the competitors tie for the leadership, have largely been neglected in the literature. Previous results establish that the most skilled agent eventually wins [19], and that average intensity up to time t is approximately $(\log t)^\alpha$, where α depends on the relative skill of the competitors [12, 27]. To the best of our knowledge, no previous work has provided rigorous characterizations for the distributions of duration and intensity of competitions in Pólya's urn models. Our work partially fills this gap for the two competitor case and sheds light on some recent approximate results [12, 27].

III.   MODELS

In this section, we formally introduce competition models for two competing agents and give precise definitions for two fundamental metrics of a competition, i.e. its duration and intensity.
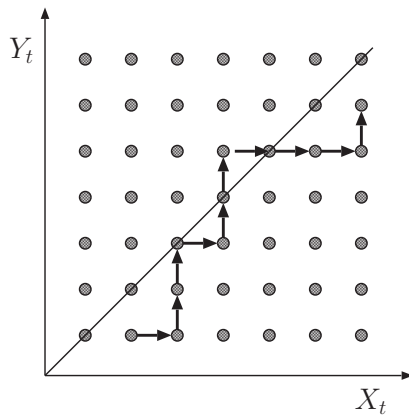
FIG. 2. State space of competition processes, with an illustration of a sample path with $x_0 = 2$, $y_0 = 1$, and three ties at time $t = 3, 5, 7$.

A. General Setup and Metrics

Let X and Y denote the two agents that engage in the competition. Each agent is associated with a positive *fitness* value that reflects its intrinsic competitiveness or skill level. Let f_X and f_Y denote the fitness of X and Y, respectively, and $r = f_X/f_Y$ the fitness ratio. Without loss of generality, we assume that $f_X \geq f_Y$ and hence $r \geq 1$.

The resource that the agents compete for will be generically referred to as wealth, which is measured in discrete units. The competition starts at time $t = 0$ with agents X and Y having x_0 and y_0 units of initial wealth, respectively. We consider a discrete-time process. At each time step, one unit of wealth is added to the system and given to either X or Y. Denote by X_t and Y_t the respective cumulative wealth of X and Y at time t. The complete history of the competition $\{(X_t, Y_t)\}_{t=0}^{\infty}$ then forms a discrete-time discrete-space stochastic process. The state space S is the first quadrant of the integral lattice (see Figure 2),

$$S = \{(x, y) \in \mathbb{Z}^2 : x \geq 1, y \geq 1\}.$$

The initial condition is $(X_0, Y_0) = (x_0, y_0)$. How the process evolves over time is defined by specific competition models, of which the CA competition model to be introduced in Section III B is an example.

We now make the notions of duration and intensity of competitions more precise by defining them through events of wealth ties. Given a competition process $\{(X_t, Y_t)\}_{t=0}^{\infty}$, we say that a *tie* occurs at time t if $X_t = Y_t$. Figure 2 shows three ties at times $t = 3, 5, 7$.

The *duration T* of a competition is defined to be the time of the last tie, i.e.,

$$T = \sup\{t \geq 0 : X_t = Y_t\}.$$

When there is no tie, we follow the standard convention that $T = \sup \emptyset = -\infty$. The competition ends at time T in the sense that one of the agents takes the lead and never lose it again after T.

The *intensity N_t* of a competition until time t is the number of ties that occur by time t, i.e.,

$$N_t = \sum_{i=0}^{t} \mathbf{1}\{X_i = Y_i\},$$

where $\mathbf{1}\{A\}$ is the indicator of event A. The *intensity N* of a competition is the total number of ties throughout the competition, i.e., $N = \lim_{t \to \infty} N_t$. This measures the intensity of the competition in the sense that it counts the number of potential changes in leadership. Note that $T < +\infty$ if and only if $N < +\infty$.

B. CA Competition Model

In the CA competition model, the unit of wealth introduced at time $t + 1$ is given to X with probability

$$p_{X,t} = \frac{f_X X_t}{f_X X_t + f_Y Y_t} = \frac{r X_t}{r X_t + Y_t};$$

otherwise it is given to Y. Note that the transition probability $p_{X,t}$ embodies both fitness and CA effects (externalities).

More formally, in the CA competition model, the complete history $\{(X_t, Y_t)\}_{t=0}^{\infty}$ forms a discrete-time Markov chain with stationary transition probabilities. The transition probability $\mathbb{P}[(X_{t+1}, Y_{t+1}) = (x', y') \mid (X_t, Y_t) = (x, y)]$ is given by

$$Q_{\text{CA},r}(x, y; x', y') = \begin{cases} \dfrac{rx}{rx + y}, & \text{if } (x', y') = (x + 1, y), \\ \dfrac{y}{rx + y}, & \text{if } (x', y') = (x, y + 1), \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Note that the transition probabilities are spatially inhomogeneous, i.e., they depend on the current state (x, y), which makes the analysis difficult, especially when $r > 1$.

For the purpose of comparison, a RW competition model incorporates skill and luck but not the CA effect (no externalities), where the transition probabilities are determined entirely by the fitness ratio r. In particular, the probability that agent X receives the unit of wealth introduced at any time is always given by

$$p_X = \frac{f_X}{f_X + f_Y} = \frac{r}{r + 1}.$$

Thus the RW competition model is a discrete-time Markov chain with the same state space S as the CA competition model, but with the following spatially homogeneous transition probabilities,

$$Q_{\text{RW},r}(x, y; x', y') = \begin{cases} \dfrac{r}{r + 1}, & \text{if } (x', y') = (x + 1, y), \\ \dfrac{1}{r + 1}, & \text{if } (x', y') = (x, y + 1), \\ 0, & \text{otherwise.} \end{cases}$$

The spatial homogeneity of the transition probabilities leads to a more tractable analysis. In fact, the difference process $\{X_t - Y_t\}$ is a standard biased RW with parameter $r/(r + 1)$. Thus the abundance of known results for RW [28] can be directly translated into results for RW competitions, including duration and intensity as we have defined in Section III A.

Throughout the rest of the paper, we use $\text{CA}_=$ and $\text{RW}_=$ to denote CA and RW competitions with identical fitness ($r = 1$), respectively. We use $\text{CA}_{\neq}$ and $\text{RW}_{\neq}$ to denote CA and RW competitions with distinct fitnesses ($r > 1$), respectively. Before presenting our results, we point out here some connections between the CA and RW models that are useful in our analysis. In particular, in $\text{CA}_=$, all paths connecting two given states (x_0, y_0) and (x, y) have the same probability. This is a nice property that $\text{CA}_=$ shares with RW, which enables us to leverage existing results on RW in our analysis of $\text{CA}_=$. Unfortunately, this property is lost in $\text{CA}_{\neq}$, where we resort to the Chapman-Kolmogorov equation for upper and lower bounds on the probabilities of interest. In the limiting case where X_t and Y_t are both large but comparable to each other, the connection to RW is again partially retained, a fact we also exploit in the analysis of $\text{CA}_{\neq}$.

IV.   RESULTS

In this section we present our theoretical results for duration and intensity distributions, which are also illustrated graphically and supported by extensive numerical simulations. Table I provides a summary of our main results along with prior knowledge about RW competitions from the literature. Note that $\mathbb{P}_{\langle \text{MODEL} \rangle, r}^{(x_0, y_0)}$ denotes the probability in model $\langle \text{MODEL} \rangle \in \{\text{CA}, \text{RW}\}$ with fitness ratio r and initial state (x_0, y_0). The following notations have been used in Table I and will be used throughout the rest of the paper.

- $f(x) \sim g(x)$ if and only if $\lim_{x \to \infty} f(x)/g(x) = 1$.

- $f(x) \lesssim g(x)$ if and only if $\limsup_{x \to \infty} f(x)/g(x) \leq 1$.

- $f(x) \gtrsim g(x)$ if and only if $\liminf_{x \to \infty} f(x)/g(x) \geq 1$.

All proofs are relegated to Appendix A.

metric	$\langle\text{Model}\rangle$	$r = 1$	$r > 1$
duration T : $\mathbb{P}^{(x_0,y_0)}_{\langle\text{Model}\rangle,r}[T \geq t]$	CA	$\sim t^{-1/2}$	$\gtrsim t^{-(r-1)x_0}$ $\lesssim t^{-(r-1)(x_0-\frac{1}{r})}$
	RW	1	$\leq \left[\frac{4r}{(r+1)^2}\right]^t$
intensity N : $\mathbb{P}^{(x_0,y_0)}_{\langle\text{Model}\rangle,r}[N \geq n]$	CA	$\sim n^{-1}$	$\leq \left(\frac{2}{r+1}\right)^{n-1}$
	RW	1	$\left(\frac{2}{r+1}\right)^{n-1}$

TABLE I. Tail distributions for duration and intensity of competitions in both RW and CA models. Multiplicative constants are omitted in all expressions involving t and n. The RW statistics can be found in most textbooks on the topic, e.g. [28, pp. 113,116].

A. Competition Duration

As shown in Table I, $\text{RW}_=$ competitions never end, i.e., $\mathbb{P}^{(x_0,y_0)}_{\text{RW},1}[T = \infty] = 1$, while $\text{RW}_{\neq}$ competitions are generally very short, whose durations exhibit exponential tails. The story for CA competitions is drastically different. The introduction of CA guarantees that a competition always ends, i.e. $\mathbb{P}^{(x_0,y_0)}_{\text{CA},r}[T < \infty] = 1$, even when the two agents are equally fit, which is in sharp contrast to endless $\text{RW}_=$ competitions. On the other hand, CA fundamentally *increases* the chance of having a long-lasting competition between unequally fit agents, as the duration of $\text{CA}_{\neq}$ always has a power-law distribution, in contrast to a sub-exponential distribution for $\text{RW}_{\neq}$. Thus, cumulative advantage does *not* always make competitions shorter as one might expect.

1. Equal Fitness Case: $CA_=$

The following theorem shows that the duration T for $\text{CA}_=$ is heavy-tailed with an asymptotic power-law distribution.

Theorem 1. *The duration of a $CA_=$ competition has the following asymptotic tail distribution,*

$$\mathbb{P}^{(x_0,y_0)}_{\text{CA},1}[T \geq t] \sim \frac{1}{2^{x_0+y_0-5/2}\sqrt{\pi}B(x_0,y_0)}t^{-1/2}, \tag{2}$$

where $B(x,y) = \int_0^1 s^{x-1}(1-s)^{y-1}ds$ is the beta function.

It follows from (2) that

$$\mathbb{P}^{(x_0,y_0)}_{\text{CA},1}[T < \infty] = 1 - \lim_{t\to\infty}\mathbb{P}^{(x_0,y_0)}_{\text{CA},1}[T \geq t] = 1,$$

i.e. the duration of $\text{CA}_=$ is almost surely finite.

Note, however, that the power-law exponent is always $-1/2$, independent of the initial wealth x_0 and y_0. Consequently, although the duration of $\text{CA}_=$ is finite rather than infinite as in $\text{RW}_=$, the expected duration is still infinite, even if x_0 is significantly larger than y_0 or vice versa.

On the other hand, the initial wealth (x_0,y_0) does affect the location of the distribution. Figure 3 shows the duration distributions from simulations for various values of initial wealth, with the asymptotes in Eq. (2) superimposed. Each simulation curve is the average of 10^5 independent runs for 10^7 time steps each. All curves are truncated at $t = 10^6$, since the empirical distributions will drop down sharply and become inaccurate as t approaches the cutoff time in simulations. Similar truncations will be applied to later plots without further mention. Note the good agreement between theory and simulation in the tails in Figure 3. When both x_0 and y_0 increase but are kept equal, the distribution curve shifts upwards, which means the competition lasts longer. When the initial wealth of only one agent (y_0 here) increases, the distribution curve shifts downwards, which means the competition is shorter.

2. Different Fitness Case: $CA_{\neq}$

The next theorem shows that the tail distribution of the duration T for $\text{CA}_{\neq}$ is asymptotically bounded by power laws from both below and above.
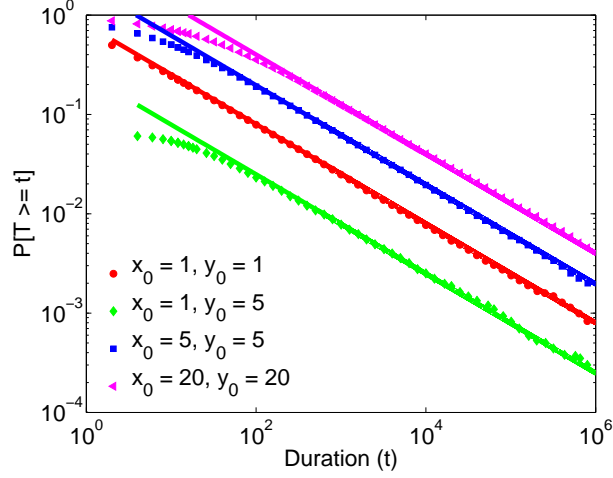
FIG. 3. Tail distribution for duration of CA$_=$ with various (x_0, y_0). The dots are simulation results. The solid lines are the asymptotes in Eq. (2).

Theorem 2. *The tail distribution of the duration of a CA$_{\neq}$ competition has the following asymptotic bounds,*

$$\varphi_1 \, t^{-(r-1)x_0} \lesssim \mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[T \geq t] \lesssim \varphi_2 \, t^{-(r-1)(x_0-1/r)}, \tag{3}$$

where

$$\varphi_1 = \frac{\Gamma(rx_0 + y_0)}{(r+1)x_0 2^{x_0+y_0-1}\Gamma(x_0)\Gamma(y_0)}, \tag{4}$$

and

$$\varphi_2 = \frac{2^{(r-1)(x_0-r^{-1})}\Gamma(r^{-1})\Gamma(rx_0 + y_0)}{(r+1)(x_0 - r^{-1})\Gamma(x_0)\Gamma(y_0)}, \tag{5}$$

where $\Gamma(x) = \int_0^\infty s^{x-1}e^{-s}ds$ is the gamma function.

It follows from the lower bound that $\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[T < \infty] = 1$ for $r > 1$, i.e. the duration for CA$_{\neq}$ is almost surely finite as is for CA$_=$. The constants φ_1 and φ_2 are very loose, so the bounds are best interpreted as bounds on the tail exponent.

Note that the power-law exponents in the upper and lower bounds depend on x_0 but not on y_0, and they differ only by $1 - 1/r < 1$. In this sense, the shape of the distribution at large t is largely determined by the fitness ratio and the initial wealth of the fitter agent, while the initial wealth of the less fit plays a much weaker role. This is illustrated in Figure 4, which shows the duration distributions from simulations for $r = 1.2$ and various values of (x_0, y_0), alongside the lower bounds from Eq. (3) that are shifted closer to the simulation results for easier comparison of the slopes. Each simulation curve is the average of 10^5 independent runs for 10^9 time steps each.

Figure 4a shows how the slopes of the distribution curves, which correspond to the power-law exponents, depend critically on x_0. The impact of x_0 is two-fold. As x_0 increases, the distribution curve becomes more tilted as predicted by the bounds. At the same time, it also shifts downwards. Both changes mean that the competition tends to be shorter.

Figure 4b shows the impact of changing both x_0 and y_0. When x_0 is fixed, increasing y_0 only results in a slight decrease in the absolute value of the slope, in agreement with Eq. (3). The distribution curve shifts upwards, which means the competition tends to last longer. When both x_0 and y_0 increase, the situation becomes more intricate. The curve may shift upwards while bending down faster in the tail, which could possibly lead to a crossover in the old and new curves, as is the case of going from $(x_0, y_0) = (1, 1)$ to $(x_0, y_0) = (3, 3)$. In this case, the new competition is more likely to have a medium long duration.

3. Struggle-of-the-Fittest Phenomenon

Now we look at the impact of fitness ratio r on duration. Contrasting Eqs. (2) and (3) leads to an interesting observation. Departing from CA$_=$ by slightly increasing the fitness ratio r from 1 to $1 + \varepsilon$, where ε is close to 0,
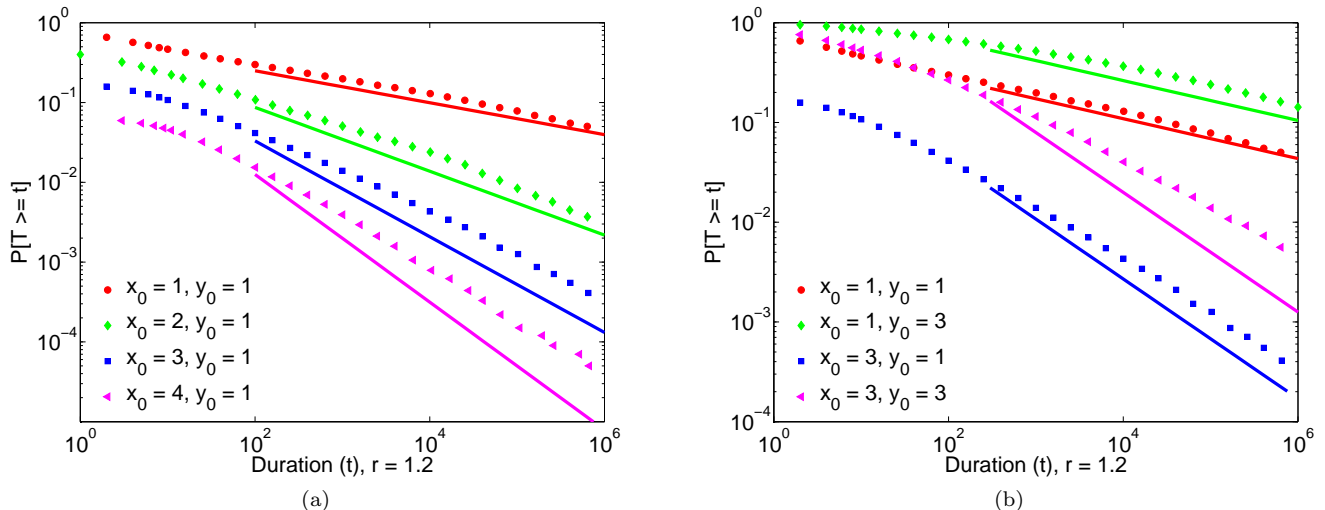
FIG. 4. Tail distribution for duration of $CA_{\neq}$ with $r = 1.2$ and various (x_0, y_0). The dots are simulation results. The solid lines are the asymptotic lower bound in Eq. (3) but shifted closer to the simulation results for easier visual comparison of the slopes.

precipitates a *significant increase* in the probability of long-lasting competitions, as manifested in the discontinuous jump in the power-law exponents from $-1/2$ in Eq. (2) to $-\varepsilon x_0 \approx 0$ in Eq. (3). This is opposite to what happens in RW competitions, where a slight increase in fitness departing from $RW_=$ to $RW_{\neq}$ transforms the competition from one that never ends to one with a geometrically distributed duration. The lower bound in Eq. (3) shows that $CA_{\neq}$ with $r < 1 + (2x_0)^{-1}$ is more likely to have long-lasting competitions than $CA_=$, despite the fact that the fitter agent is bound to become the ultimate winner. We refer to the phenomenon that the fitter agent takes an extremely long time to win as "struggle of the fittest".

Figure 5 shows the duration of simulated CA competitions for various fitness ratios r. Each simulation curve is the average of 10^5 independent runs for 10^9 time steps each. Note how the distribution of duration jumps upward from the curve for $CA_=$ to the curve for $CA_{\neq}$ with $r = 1.1$. It also shows how the curves for $CA_{\neq}$ become more and more tilted as r increases, being roughly parallel to the $CA_=$ curve at $r = 1 + (2x_0)^{-1} = 1.5$.
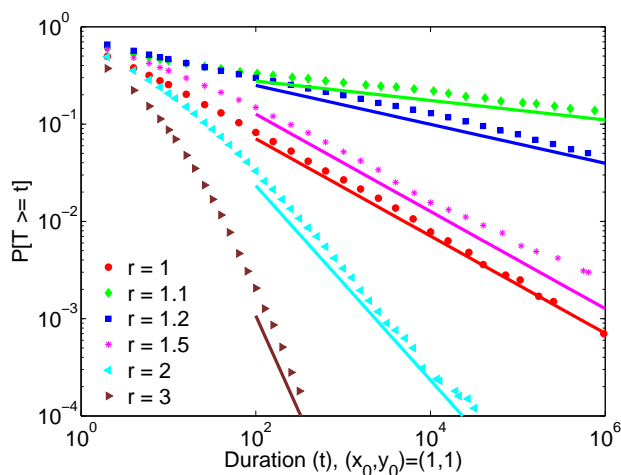

FIG. 5. Tail distribution for duration of CA with various r. The dots are simulation results. For $r = 1$, the solid line is the asymptote in Eq. (2). For $r > 1$, the solid lines are the lower bound in Eq. (3) but shifted as in Figure 4.

B. Competition Intensity

Given that CA competitions are long-lasting, one might expect them also to be intense, i.e., exhibit many ties ($X_t = Y_t$). As we will see in this section, this intuition is appropriate for $CA_=$ but not for $CA_{\neq}$.

1. Equal Fitness Case: $CA_=$

The following theorem shows that the intensity N of $CA_=$ is heavy-tailed with an asymptotic power-law distribution.

Theorem 3. *The intensity of a $CA_=$ competition has the following asymptotic tail distribution,*

$$\mathbb{P}_{CA,1}^{(x_0,y_0)}[N \geq n] \sim \frac{1}{2^{x_0+y_0-2}B(x_0,y_0)}n^{-1}, \tag{6}$$

where $B(x_0,y_0)$ is the beta function as in Eq. (2).

In this case, the intensity has infinite expectation, as does the duration. Figure 6 shows the duration distributions from simulations for various values of initial wealth, with the asymptotes in Eq. (6) superimposed. Each simulation curve is the average of 10^5 independent runs for 10^7 time steps each. We observe the same behavior as in Figure 3. When both x_0 and y_0 increase but are kept equal, the distribution curve shifts upwards, which means the competition is more intense. When the initial wealth of only one agent (y_0 here) increases, the distribution curve shifts downwards, which means the competition is less intense.

We mention in passing that if we have a finite observation time t_f, the expected intensity N_{t_f} by time t_f grows as $\log t_f$, a phenomenon observed for the related CA model in Godrèche et al. [12].
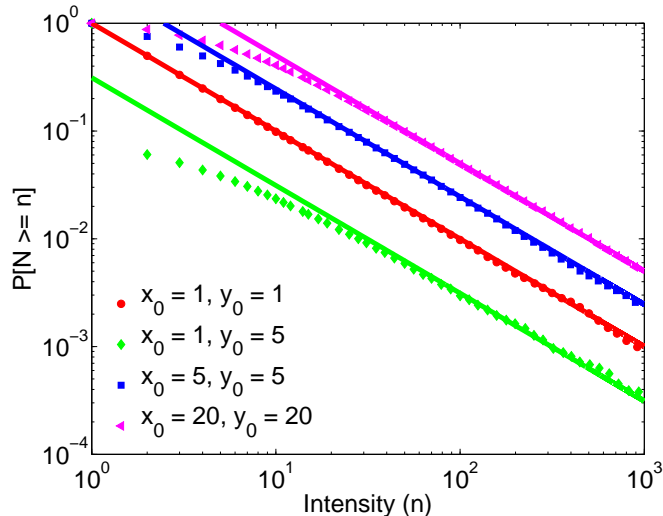

FIG. 6. Tail distribution for intensity of $CA_=$ with various (x_0,y_0). The dots are simulation results. The solid lines are the asymptotes from Eq. (6).
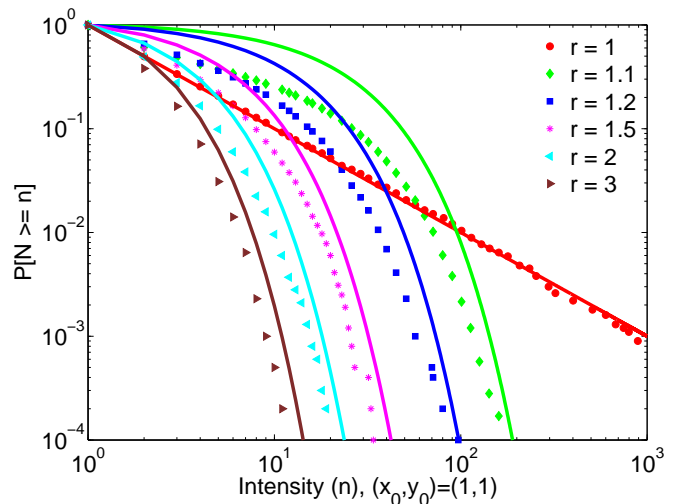
FIG. 7. Tail distribution for intensity of CA with various r. The dots are simulation results. The solid lines are the upper bounds from Eq. (7) for $r > 1$, and the asymptote in Eq. (6) for $r = 1$.

2. Different Fitness Case: $CA_{\neq}$

In sharp contrast, $CA_{\neq}$ competitions are not intense despite their long durations. In fact their intensities are surprisingly mild, bounded above by a geometric distribution, as shown in the next theorem.

Theorem 4. *The tail distribution of the intensity of a $CA_{\neq}$ competition has the following upper bound,*

$$\mathbb{P}_{CA,r}^{(x_0,y_0)}[N \geq n] \leq C\left(\frac{2}{1+r}\right)^{n-1}, \tag{7}$$

with

$$C = \begin{cases} 1, & x_0 \leq y_0, \\ \frac{(y_0)_{x_0-y_0}}{(rx_0+y_0)_{x_0-y_0}} \left(1 + \frac{1}{r}\right)^{x_0-y_0}, & x_0 > y_0, \end{cases}$$

where $(x)_k = \prod_{i=0}^{k-1}(x+i)$ *is the Pochhammer symbol.*

Note that the expectation and all higher moments of N are finite. Therefore, the intensity of a CA competition changes dramatically when the fitnesses of the two parties become unequal, the distribution shifting from a power-law tail to an exponential tail. This is illustrated in Figure 7, where each simulation curve is the average of 10^5 independent runs for 10^9 time steps each. An important observation is that both CA$_{\neq}$ and RW$_{\neq}$ competitions have intensities that are upper bounded by identical exponential tails (see Table I), while exhibiting fundamentally different durations.

Why are CA$_{\neq}$ competitions simultaneously not intense and long-lasting? The answer resides in the probability of Y being the eventual winner. In CA$_{=}$ competitions, Y wins with probability $y_0/(x_0 + y_0)$, while in CA$_{\neq}$ competitions Y (the less fit) never wins. However, for small values of r, especially for those very close to one, the dynamics in the initial stages of the competition closely follows that of CA$_{=}$. Thus there is a non-negligible chance that Y takes a significant lead, with the CA effect helping it uphold the lead for a long period of time over which there is no tie. Eventually, however, fitness effect outweighs the CA effect, and X catches up with Y. By then they both have large accumulated wealth, which makes CA$_{\neq}$ behave like RW$_{\neq}$ in the vicinity of $X = Y$, allowing X to quickly establish a lead ahead of Y. At this final stage both fitness and CA effects work in favor of X, and Y stands little chance in taking the lead again. To summarize, the less fit agent has a non-negligible probability of taking an early lead which can last for a very long time due to the CA effect, but it will ultimately surrender the lead to the fitter agent and never lead again, a phenomenon that we call "delusion of the weakest", which is the flip-side of "struggle of the fittest".

Figure 8 illustrates this observation by showing sample paths for different values of r, all generated using the same sequence of random bits. Note that for $r = 1$ (identical fitness), Y wins quickly, whereas for $r = 1.1$ the fitter agent X, having trailed behind for a long time, eventually takes over after 69,426 time steps. Finally, for both $r = 1.2$ and $r = 1.5$ agent X has no trouble quickly winning the competition. These sample paths showcase the long struggle of the "slightly" fitter agent in competitions with CA effects.
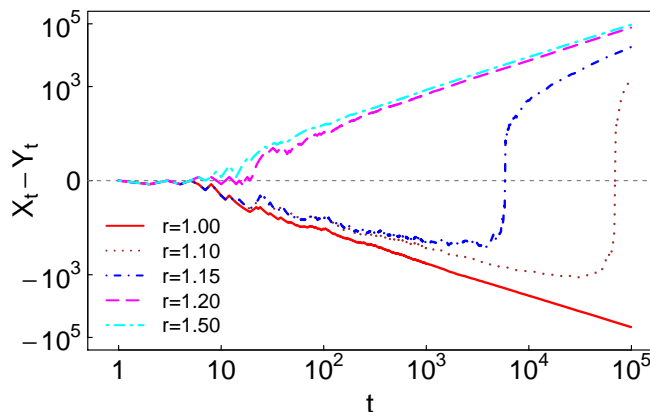

FIG. 8. "Delusion of the weakest": sample paths for different values of r ($x_0 = y_0 = 1$), all generated using the same sequence of random bits

C. Interplay of Duration and Intensity

In this section, we study the relationship between duration and intensity. Note that duration gives a natural upper bound $N \leq T/2$ for intensity, i.e., the number of ties is at most half of the duration in any competition. In CA$_{=}$, duration and intensity are strongly and positively correlated. In fact, a tie at time t increases the probability of having another tie at a time greater than t. More precisely, [25] shows that for CA$_{=}$,

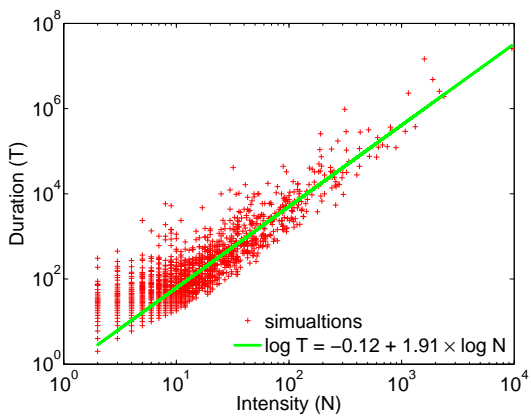$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[T > t | X_t = Y_t] \simeq 1 - \frac{1}{\sqrt{\pi X_t}}. \tag{8}$$
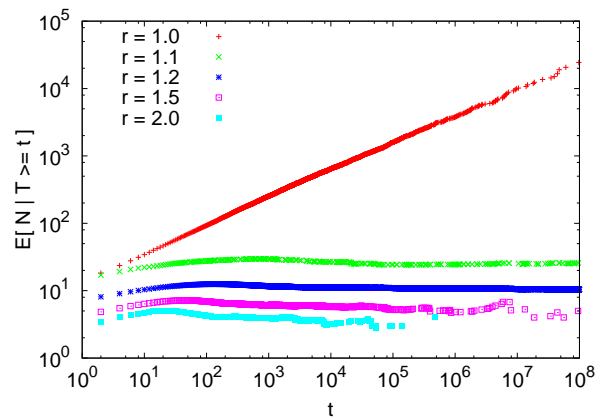
FIG. 9. Scatter plot of duration vs. intensity for $CA_=$.

FIG. 10. Conditional average intensity of competitions conditioned on their duration being at least t, namely $\mathbb{E}_{CA,r}^{(x_0,y_0)}[N \mid T \geq t]$.

Since $X_t \sim t/2$ at a tie, Eq. (8) implies that the later a tie occurs, the more likely another tie will occur, intuitively explaining why long-lasting competitions are also intense in this case.

Figure 9 shows a scatter-plot of duration versus intensity from 10^4 independent runs of $CA_=$ competitions with $x_0 = y_0 = 1$, each simulated for 10^9 time steps. This unveils a strong positive linear correlation between the two statistics in log-log scale (sample Pearson correlation coefficient of 0.94).

Interestingly, $CA_{\neq}$ shows a different behavior, since even long-lasting competitions exhibit only a small number of ties. Figure 10 shows simulation results for conditional average intensities of competitions with $x_0 = y_0 = 1$ and different fitness ratios r, conditioned on the duration being at least t. Each simulation curve is obtained from 10^4 independent runs for 10^9 time steps each. Note that for $r = 1$, the conditional average intensity increases linearly with t, but for $r > 1$, it stabilizes as t increases. Again, we observe a sharp transition as we move from identical to distinct fitnesses, this time in the correlation between intensity and duration.

V.   DISCUSSION AND CONCLUSION

As supported by various empirical studies over the last century, real world competitions for resource accumulation seem to be subject to cumulative advantage effects, at least to some extent. Because of such findings and the fact that CA generally leads to large inequalities in resource distribution, understanding the role of skill and luck in competition dynamics becomes a pressing issue both in theory and practice. Indeed, recent empirical [2, 29, 30] and theoretical [1, 12, 25, 31] studies have contributed in this direction.

However, contrary to prior theoretical works, we consider simple and classical mathematical models that capture just the essence of skill and luck competitions with and without CA effects, and investigate fundamental aspects of competition, namely duration (i.e., time until an ultimate winner emerges) and intensity (i.e., number of ties in competition). By considering simple models and simple properties we prove and illustrate fundamental theoretical results: CA effect exacerbates the role of luck - power-law tail duration emerges regardless of skill differences, and become extreme (i.e., infinite mean) when skill differences are small enough. Moreover, duration is long not necessarily because of intense competition where agents tussle aggressively for ultimate leadership. On the contrary, under CA, competitions are generally very mild, exhibiting an exponential tail. Long competitions emerge when an early stroke of luck places the less skilled in the lead, who can then, boosted by CA effects, enjoy leadership for a very long period of time. Thus, when CA is present luck sides with the less skilled.

The non-negligible probability of long-lasting competitions has far-reaching implications. In the absence of CA, it takes very little time for the fittest agent to establish dominance, so it is often reasonable to neglect the possibility of a premature burnout. Such observations are in hand with the "survival of the fittest" principle, since soon enough the more skilled will prevail. In the presence of CA, however, even agents with superior fitness may face the challenge of having to endure extremely long competitions. This challenge becomes all more real when the fitness superiority is only minimal. Will the more skilled survive the seemingly eternal inferiority during the competition? Under CA time becomes a central issue, with delusion becoming reality if the more skilled burns out during a long struggle. Thus, in the face of CA, the fittest survives only if it can persist, which prompts us to rename the principle "survival of the fittest and persistent" when considering CA competitions.

This observation may also shed light on the seemingly inherent difficulty of predicting success in real-world competitions by observing ongoing sample paths. Different empirical studies have alluded to this problem [2, 29, 30] as well as recent model driven studies [31]. A key ingredient in predicting success is the assessment of the agents' fitnesses. However, fitness estimation for an ongoing competition which is driven by a single, likely short, sample path is a hard problem. The Cramér-Rao bound [32] provides a lower bound on the mean squared error of any unbiased estimator. The Cramér-Rao bound of the fitness values estimates $\hat{f}_X$ and $\hat{f}_Y$ in the linear CA case with one step is

$$\text{MSE}(\hat{f}_h) = (J^{-1})_{hh} \, , h \in \{X, Y\} \, ,$$

where

$$J_{i,j} = \sum_{z_X = X_0}^{X_0+1} \sum_{z_Y = Y_0}^{Y_0+X_0+1-z_X} \left(\frac{(f_i^2 - f_i)z_i + f_i f_j z_j}{(f_i z_i + f_j z_j)^2} \right) \left(\frac{(f_j^2 - f_j)z_j + f_j f_i z_i}{(f_j z_j + f_i z_i)^2} \right) \left(\frac{f_i z_i}{f_i z_i + f_j z_j} \right).$$

Replacing fitness values $f_X = 1$ and $f_Y = 2$, with $X_0 = Y_0 = 1$ into the above equation we get $\text{MSE}(\hat{f}_X) \geq 105$ and $\text{MSE}(\hat{f}_Y) \geq 59$. These are large lower bounds for the error, showing that accurate unbiased estimation is not possible in this case. In particular, in light of our results which indicate that competition durations follow power laws, an accurate estimation of agents' fitnesses may require extremely long observations. In an environment where direct competitions tend to be relatively short with a competitor giving up after some time and fitness differences among agents tend to be small, estimating agents' fitnesses or even the most fit agent seems challenging.

The CA effect considered in the present work is linear, which need not be the case in general. Nonlinear CA has also been investigated in the literature. In particular, [22, 23] considered the case $p_{X,t} = \frac{rX_t^\beta}{rX_t^\beta + Y_t^\beta}$ with $r = 1$ and $\beta > 0$ (note that $\beta = 1$ corresponds to linear CA). A natural problem here is to characterize the duration and intensity of nonlinear CA competitions, where the interaction between CA and fitness becomes more complicated. Our preliminary investigation in this direction indicates that in the equal fitness case ($r = 1$) both duration and intensity retain power-law distributions, whose exponents depend on β. When agents have different fitnesses ($r > 1$) intensity seems to remain light-tailed for all values of β. In contrast, duration seems to be heavy-tailed only when $\beta \geq 1$ and become light-tailed when $\beta < 1$. A more precise and detailed characterization of non-linear CA, however, is more involved and left for future work.

[1] PL Krapivsky and S Redner. Statistics of changes in lead node in connectivity-driven networks. *Physical review letters*, 89(25):258703, 2002.

[2] Dashun Wang, Chaoming Song, and Albert-László Barabási. Quantifying long-term scientific impact. *Science*, 342(6154):127–132, 2013.

[3] Alexander M. Petersen, Woo-Sung Jung, Jae-Suk Yang, and H. Eugene Stanley. Quantitative and empirical demonstration of the matthew effect in a study of career longevity. *Proceedings of the National Academy of Sciences*, 108(1):18–23, 2011.

[4] W. B. Arthur. *Increasing Returns and Path Dependence in the Economy*. U. Michigan Press, 1994.

[5] Thomas A DiPrete and Gregory M Eirich. Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual review of sociology*, pages 271–297, 2006.

[6] Florian Eggenberger and George Pólya. Über die statistik verketteter vorgänge. *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 3(4):279–289, 1923.

[7] Eli Eisenberg and Erez Y Levanon. Preferential attachment in the protein network evolution. *Phy. Rev. Let.*, 91(13):138701, 2003.

[8] NG van Kampen. Stochastic processes in physics and chemistry. *North-Holland*, 1, 1981.

[9] Youmna Borghol, Siddharth Mitra, Sebastien Ardon, Niklas Carlsson, Derek Eager, and Anirban Mahanti. Characterizing and modelling popularity of user-generated videos. *Performance Evaluation*, 68(11):1037–1055, 2011.

[10] Flavio Figueiredo, Fabrício Benevenuto, and Jussara M Almeida. The tube over time: characterizing popularity growth of youtube videos. In *ACM international conference on Web search and data mining*, pages 745–754, 2011.

[11] D. de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292–306, 1976.

[12] C Godrèche, H Grandclaude, and J M Luck. Statistics of leaders and lead changes in growing networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(02):P02001, 2010.

[13] Vincent D. Blondel, Jean-Loup Guillaume, Julien M. Hendrickx, Cristobald de Kerchove, and Renaud Lambiotte. Local leaders in random networks. *Phys. Rev. E*, 77:036114, Mar 2008.

[14] Albert-László Barabási. Network science: Luck or reason. *Nature*, 489(7417):507–508, 2012.

[15] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[16] G. Bianconi and A.-L. Barabási. Competition and multiscaling in evolving networks. *Europh. Let.*, 54(4):436–442, 2001.

[17] R. K. Merton. The matthew effect in science. *Science*, 159:56–63, 1968.

[18] Matjaž Perc. The matthew effect in empirical data. *Journal of The Royal Society Interface*, 11(98), 2014.

[19] Hosam Mahmoud. *Pólya urn models*. CRC Press, 2008.

[20] S. Janson. Limit theorems for triangular urn schemes. *Prob. Theory Related Fields*, 134:417–452, 2005.

[21] R. Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, 4:1–79, 2007.

[22] Eleni Drinea, Alan Frieze, and Michael Mitzenmacher. Balls and bins models with feedback. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 308–315, 2002.

[23] Roberto Oliveira. Balls-in-bins processes with feedback and brownian motion. *Journal of Combinatorics, Probability and Computing*, 17(1):87–110, 2008.

[24] S. Bubeck, E. Mossel, and M. Rácz. On the influence of the seed graph in the preferential attachment model. *Network Science and Engineering, IEEE Transactions on*, PP(99):1–1, 2015.

[25] T. Antal, E. Ben-Naim, and PL Krapivsky. First-passage properties of the Pólya urn process. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(07):P07009, 2010.

[26] T.C. Wallstrom. The equalization probability of the Pólya urn. *The American Mathematical Monthly*, 119(6):516–518, 2012.

[27] C Godrèche and J M Luck. On leaders and condensates in a growing network. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(07):P07031, 2010.

[28] Barry D Hughes. *Random Walks and Random Environments*, volume 1: Random Walks. Oxford University Press, 1995.

[29] Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856, 2006.

[30] Arnout van de Rijt, Soong Moon Kang, Michael Restivo, and Akshay Patil. Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 111(19):6934–6939, 2014.

[31] Jerker Denrell and Chengwei Liu. Top performers are not the most impressive when extreme performance indicates unreliability. *Proceedings of the National Academy of Sciences*, 109(24):9331–9336, 2012.

[32] Bruno Ribeiro, Don Towsley, Tao Ye, and Jean C Bolot. Fisher information of sampled packets: an application to flow size estimation. In *Proc. 6th ACM SIGCOMM Conf. Internet Meas.*, pages 15–26. ACM, 2006.

[33] Norman L Johnson, Adrienne W Kemp, and Samuel Kotz. *Univariate Discrete Distributions*. John Wiley & Sons, 2nd edition, 1992.

[34] Frank WJ Olver. *NIST handbook of mathematical functions*. Cambridge University Press, 2010.

[35] Willliam Feller. *An introduction to probability theory and its applications*, volume 1. John Wiley & Sons, 3rd edition, 1968.

[36] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110(2):177–245, 2004.

Appendix A: Proofs

In this appendix, we provide proofs for our main results in Section IV. We will use the following additional notations and definitions.

- Denote the transition probability from state (x_0, y_0) to state (x, y) in $t = x + y - x_0 - y_0$ steps by

$$p_r(x_0, y_0; x, y) = \mathbb{P}_{\text{CA},r}^{(x_0, y_0)}[(X_t, Y_t) = (x, y)].$$

Note that the transition probability is nonzero only for this specific t. Thus we will often omit to mention t explicitly hereafter and assume that the appropriate t has been chosen.

- Let τ_n be the time of the n-th tie, which can be defined recursively by $\tau_0 = -\infty$ and

$$\tau_n = \inf\{t > \tau_{n-1} : X_t = Y_t\}, \ n \geq 1.$$

Note that $T = \tau_N$.

- Denote by $q_r(x_0, y_0)$ the probability of having no tie after leaving state (x_0, y_0), i.e.,

$$q_r(x_0, y_0) = \mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[X_t \neq Y_t, t \geq 1].$$

Note that $q_r(x_0, x_0) = \mathbb{P}_{\text{CA},r}^{(x_0,x_0)}[\tau_2 = \infty]$ and $q_r(x_0, y_0) = \mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\tau_1 = \infty]$ for $x_0 \neq y_0$.

- Denote by $A_{n,t}(x, y)$ the set of paths that start from (x, y) at time 0 and end with the n-th tie at time t, i.e., $\tau_n = t$.

1.   Proof of Theorem 1

Note that the $\text{CA}_=$ model is the standard Pólya urn model. The proof of Theorem 1 combines known results for this model. Starting from the initial state (x_0, y_0), X_t has a beta-binomial distribution with parameters x_0 and y_0 [33]. Note that the event $X_t = Y_t$ occurs only if $t = |x_0 - y_0| + 2k$ for some integer $k \geq 0$. For such t, $X_t = Y_t$ if and only if $X_t = z_0 + k$, where $z_0 = \max\{x_0, y_0\}$. By Eq. (6.27) of [33],

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[X_t = Y_t] = \mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[X_t = z_0 + k] = \frac{B(z_0 + k, z_0 + k)}{B(x_0, y_0)}\binom{t}{k}. \tag{A1}$$

Recall that $q_1(x, y)$ is the probability of having no tie after leaving state (x, y). Thus

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[T = t] = \mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[X_t = Y_t] \cdot q_1(z_0 + k, z_0 + k), \tag{A2}$$

where the second factor on the right-hand side is the probability of having no tie after t.

Recall that the exit probability $E(x, y)$ in [25] is the probability of ever having a tie starting from (x, y), including the initial state (x, y). Thus for $x \neq y$, $q_1(x, y)$ is related to $E(x, y)$ by

$$q_1(x, y) = 1 - E(x, y).$$

Using Eq. (22) of [25] for $E(x, y)$, we obtain

$$q_1(x + 1, x) = q_1(x, x + 1) = \frac{\Gamma(x + 1/2)}{\Gamma(x + 1)\Gamma(1/2)}.$$

However, $q_1(x, x) \neq E(x, x) = 1$. By considering the one-step transition from (x, x) to $(x + 1, x)$ or $(x, x + 1)$, we obtain

$$q_1(x, x) = \frac{1}{2}q_1(x + 1, x) + \frac{1}{2}q_1(x, x + 1) = \frac{\Gamma(x + 1/2)}{\Gamma(x + 1)\Gamma(1/2)}.$$

Eliminating $\Gamma(x + 1/2)$ by the identity

$$\Gamma(2x) = \pi^{-1/2}2^{2x-1}\Gamma(x)\Gamma(x + 1/2)$$

in [34, Eq. (5.5.5)], and using $\Gamma(x + 1) = x\Gamma(x)$ and $\Gamma(1/2) = \sqrt{\pi}$, we obtain

$$q_1(x, x) = \frac{\Gamma(2x)}{x2^{2x-1}\Gamma(x)\Gamma(x)} = \frac{1}{x2^{2x-1}B(x, x)}. \tag{A3}$$

Substitution of Eqs. (A1) and (A3) into Eq. (A2) yields

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[T = t] = \frac{1}{B(x_0, y_0)} \cdot \frac{1}{(z_0 + k)2^{2k+2z_0-1}}\binom{t}{k}.$$

For $t = |x_0 - y_0| + 2k$, Stirling's formula yields

$$\binom{t}{k} \sim \sqrt{\frac{2}{\pi}}t^{-1/2}2^t,$$

and hence

$$\mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[T=t] \sim \frac{1}{2^{x_0+y_0-5/2}\sqrt{\pi}B(x_0,y_0)}t^{-3/2}. \tag{A4}$$

It is well-known that as $t \to \infty$, $X_t/(X_t+Y_t)$ converges almost surely to a beta random variable V. It follows that $|X_t - Y_t|/(X_t+Y_t) \to |2V-1|$. Thus, for $V \neq 1/2$, which holds almost surely, we have $|X_t - Y_t|/(X_t+Y_t) > 0$ for all large enough t. Therefore, $\mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[T=\infty]=0$. Summing over t in Eq. (A4), we obtain as $t \to \infty$,

$$\mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[T\geq t] = \sum_{t'=t}^{\infty} \mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[T=t'] \sim \frac{1}{2}\sum_{s=t}^{\infty}\frac{1}{2^{x_0+y_0-5/2}\sqrt{\pi}B(x_0,y_0)}s^{-3/2}$$

$$\sim \frac{1}{2}\int_t^{\infty}\frac{1}{2^{x_0+y_0-5/2}\sqrt{\pi}B(x_0,y_0)}s^{-3/2}ds = \frac{1}{2^{x_0+y_0-5/2}\sqrt{\pi}B(x_0,y_0)}t^{-1/2},$$

where we have used the fact that half of the terms are zero in the second step, and $\sum_{s=t}^{\infty}s^{-a} \sim \int_t^{\infty}s^{-a}ds$ in the third step. This completes the proof of Theorem 1.

2. Proof of Theorem 2

Similar to Eq. (A2), we have

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T=t] = p_r(x_0,y_0;z_0+k,z_0+k) \cdot q_r(z_0+k,z_0+k). \tag{A5}$$

Thus the proof here amounts to finding expressions for both $p_r(x_0,y_0;z_0+k,z_0+k)$ and $q_r(z_0+k,z_0+k)$ in $\mathrm{CA}_{\neq}$. We break the proof into three lemmas.

Lemma 1.

$$p_r(x_0,y_0;x_0+k,y_0+h) \geq \frac{(x_0)_k(y_0)_h}{(rx_0+y_0)_{k+h}}\binom{k+h}{k}, \tag{A6}$$

for all $k \geq 0, h \geq 0$.

Lemma 2.

$$p_r(x_0,y_0;x_0+k,y_0+h) \leq \frac{(x_0)_k(y_0)_h}{(r^{-1})_k(rx_0+y_0)_h}, \tag{A7}$$

for all $k \geq 0, h \geq 0$.

Lemma 3. *For $r > 1$,*

$$q_r(x,x) \to \frac{r-1}{r+1}, \tag{A8}$$

as $x \to \infty$.

Before proving these lemmas, we first use them to prove Theorem 2.

Proof of Theorem 2. By Lemma 1, we have

$$p_r(x_0,y_0;z_0+k,z_0+k) \geq \frac{(x_0)_{k+z_0-x_0}(y_0)_{k+z_0-y_0}}{(rx_0+y_0)_{2k+2z_0-x_0-y_0}}\binom{2k+2z_0-x_0-y_0}{k+z_0-x_0}$$

$$= \frac{\Gamma(rx_0+y_0)}{\Gamma(x_0)\Gamma(y_0)} \cdot \frac{\Gamma(k+z_0)\Gamma(k+z_0)}{\Gamma(k+z_0-x_0+1)\Gamma(k+z_0-y_0+1)} \cdot \frac{\Gamma(2k+2z_0-x_0-y_0+1)}{\Gamma(2k+2z_0+(r-1)x_0)}$$

Using the relation $\Gamma(k+a)/\Gamma(k+b) \sim k^{a-b}$ as $k \to \infty$, we obtain

$$p_r(x_0,y_0;z_0+k,z_0+k) \gtrsim \frac{\Gamma(rx_0+y_0)}{2^{x_0+y_0-2}\Gamma(x_0)\Gamma(y_0)}(2k)^{-(r-1)x_0-1} = 2(r+1)x_0\varphi_1(2k)^{-(r-1)x_0-1},$$

where φ_1 is given by Eq. (4). Application of this asymptotic bound and Lemma 3 to Eq. (A5) yields

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T = t] \gtrsim 2(r-1)x_0\varphi_1 t^{-(r-1)x_0-1}. \tag{A9}$$

Note that $\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T = +\infty] = \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[N = +\infty] = 0$, where the second equality will follow from Theorem 4, so we will not provide a separate proof here. Summing over t in Eq. (A9) and noting that half of the terms are zero, we obtain as $t \to \infty$,

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T \geq t] = \sum_{t'=t}^{\infty} \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T = t'] \gtrsim \int_t^{\infty} (r-1)x_0\varphi_1 s^{-(r-1)x_0-1} ds = \varphi_1 t^{-(r-1)x_0},$$

establishing the lower bound.

In a similar way, Lemma 2 yields

$$p_r(x_0,y_0; z_0 + k, z_0 + k) \lesssim 2(r+1)(x_0 - r^{-1})\varphi_2(2k)^{-(r-1)(x_0-r^{-1})-1},$$

where φ_2 is given by Eq. (5). Application of this asymptotic bound and Lemma 3 to Eq. (A5) yields

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T = t] \lesssim 2(r-1)(x_0 - r^{-1})\varphi_2 t^{-(r-1)(x_0-r^{-1})-1},$$

and

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T \geq t] = \sum_{t'=t}^{\infty} \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[T = t'] \lesssim \varphi_2 t^{-(r-1)(x_0-r^{-1})},$$

establishing the upper bound. $\qquad\square$

Now we prove the lemmas. Recall that the transition probability $p(x_0, y_0; x, y)$ of going from (x_0, y_0) to (x, y) satisfies the following recursion (Chapman-Kolmogorov equation),

$$p_r(x_0, y_0; x, y) = \frac{r(x-1)}{r(x-1) + y}p_r(x_0, y_0; x-1, y) + \frac{y-1}{rx + y - 1}p_r(x_0, y_0; x, y-1), \tag{A10}$$

for $x \geq x_0$, $y \geq y_0$ and $x + y \geq x_0 + y_0 + 1$, with the boundary condition $p_r(x_0, y_0; x, y) = 0$ for $x < x_0$ or $y < y_0$. Note that we have replaced the one-step transition probabilities $Q_{\mathrm{CA},r}(x-1, y; x, y)$ and $Q_{\mathrm{CA},r}(x, y-1; x, y)$ by the expressions in Eq. (1).

Proof of Lemma 1. We will use the short-hand notation $p(k, h)$ for $p_r(x_0, y_0; x_0 + k, y_0 + h)$, and $\psi(k, h)$ for the right-hand side of Eq. (A6). We first prove the boundary case for $k = 0$. By Eq. (A10), for $h \geq 1$,

$$p(0, h) = \frac{y_0 + h - 1}{rx_0 + y_0 + h - 1}p(0, h-1),$$

which is a simple recursion in h and can be expanded to yield

$$p(0, h) = \frac{(y_0)_h}{(rx_0 + y_0)_h}p(0, 0) = \frac{(y_0)_h}{(rx_0 + y_0)_h} = \psi(0, h),$$

which yields Eq. (A6) for $k = 0$ and $h \geq 1$. Here we have used $p(0, 0) = p_r(x_0, y_0; x_0, y_0) = 1$.

Similarly, for the other boundary case $h = 0, k \geq 1$, we have

$$p(k, 0) = \frac{(x_0)_k}{(x_0 + r^{-1}y_0)_k}p(0, 0) = \frac{(x_0)_k}{(x_0 + r^{-1}y_0)_k} \geq \frac{(x_0)_k}{(rx_0 + y_0)_k} = \psi(k, 0),$$

where the last inequality is because $(x)_k$ increases with x, and $x_0 + r^{-1}y_0 \leq rx_0 + y_0$.

For the general case, we use induction on $k + h$. The base case $k + h = 1$ is already proven, since either $k = 0$ or $h = 0$ when $k + h = 1$. Assume Eq. (A6) holds for $k + h = m \geq 1$. Consider $k + h = m + 1$. We can also assume $k \geq 1$ and $h \geq 1$, since we have proven the boundary cases for $k = 0$ or $h = 0$. The recursion in Eq. (A10) yields

$$p(k, h) = \frac{r(x_0 + k - 1)}{rk + h + c_0 - r}p(k-1, h) + \frac{y_0 + h - 1}{rk + h + c_0 - 1}p(k, h-1)$$

$$\geq \frac{r(x_0 + k - 1)}{rk + h + c_0 - 1}p(k - 1, h) + \frac{y_0 + h - 1}{rk + h + c_0 - 1}p(k, h - 1),$$

where $c_0 = rx_0 + y_0$.

Applying the induction hypothesis $p(k - 1, h) \geq \psi(k - 1, h)$ and $p(k, h - 1) \geq \psi(k, h - 1)$ to the above inequality yields

$$p(k, h) \geq \frac{r(x_0 + k - 1)}{rk + h + c_0 - 1}\psi(k - 1, h) + \frac{y_0 + h - 1}{rk + h + c_0 - 1}\psi(k, h - 1) = \frac{(rk + h)(k + h + c_0 - 1)}{(k + h)(rk + h + c_0 - 1)}\psi(k, h),$$

where in the last step we have used

$$\psi(k - 1, h) = \frac{k}{k + h} \cdot \frac{k + h + c_0 - 1}{x_0 + k - 1}\psi(k, h),$$

and

$$\psi(k, h - 1) = \frac{h}{k + h} \cdot \frac{k + h + c_0 - 1}{y_0 + h - 1}\psi(k, h).$$

To complete the proof, it suffices to show that

$$\frac{(rk + h)(k + h + c_0 - 1)}{(k + h)(rk + h + c_0 - 1)} \geq 1,$$

but this is equivalent to $r \geq 1$, which is true by assumption. $\qquad\square$

Proof of Lemma 2. The proof of Lemma 2 follows the same line of reasoning as that used to prove Lemma 1. The boundary cases can be verified directly. We only outline the induction step here. Applying Eq. (A7) to the right-hand side of Eq. (A10) yields

$$p(k, h) \leq \frac{r(x_0 + k - 1)}{rk + h + c_0 - r} \frac{(x_0)_{k-1}(y_0)_h}{(r^{-1})_{k-1}(c_0)_h} + \frac{y_0 + h - 1}{rk + h + c_0 - 1} \frac{(x_0)_k(y_0)_{h-1}}{(r^{-1})_k(c_0)_{h-1}}$$
$$= \left[\frac{r(r^{-1} + k - 1)}{rk + h + c_0 - r} + \frac{c_0 + h - 1}{rk + h + c_0 - 1}\right] \frac{(x_0)_k(y_0)_h}{(r^{-1})_k(c_0)_h}.$$

Note that

$$\frac{r(r^{-1} + k - 1)}{rk + h + c_0 - r} + \frac{c_0 + h - 1}{rk + h + c_0 - 1} \leq \frac{r(r^{-1} + k - 1)}{rk + h + c_0 - r} + \frac{c_0 + h - 1}{rk + h + c_0 - r} = 1,$$

which completes the induction. $\qquad\square$

Proof of Lemma 3. Recall that $A_{n,2k}(x, x)$ is the set of paths that start from (x, x) at time 0 and end with the n-th tie at time $2k$, i.e., $\tau_n = 2k$. Let $A_{n,2k} = A_{n,2k}(0, 0)$. Note that the paths in $A_{n,2k}(x, x)$ are exactly the paths in $A_{n,2k}$ translated by (x, x). Let $\tilde{\pi} \in A_{n,2k}$ and its state at time t be $\tilde{\pi}_t = (\tilde{x}_t, \tilde{y}_t)$. The translation of $\tilde{\pi}$ by (x, x), denoted $x + \tilde{\pi}$, is a path in $A_{n,2k}(x, x)$, whose probability in the CA model is given by

$$\mathbb{P}_{\mathrm{CA},r}^{(x,x)}[x + \tilde{\pi}] = \prod_{j=0}^{2k-1}\left(\frac{r(x + \tilde{x}_j)}{r(x + \tilde{x}_j) + (x + \tilde{y}_j)}\right)^{\tilde{x}_{j+1} - \tilde{x}_j} \cdot \prod_{j=0}^{2k-1}\left(\frac{(x + \tilde{y}_j)}{r(x + \tilde{x}_j) + (x + \tilde{y}_j)}\right)^{\tilde{y}_{j+1} - \tilde{y}_j}.$$

For fixed k and $\tilde{\pi}$, as $x \to \infty$, $\mathbb{P}_{\mathrm{CA},r}^{(x,x)}[x + \tilde{\pi}]$ converges to

$$\prod_{j=0}^{2k-1}\left(\frac{r}{r + 1}\right)^{\tilde{x}_{j+1} - \tilde{x}_j}\left(\frac{1}{r + 1}\right)^{\tilde{y}_{j+1} - \tilde{y}_j} = \mathbb{P}_{\mathrm{RW},r}^{(0,0)}[\tilde{\pi}],$$

which corresponds to the probability of the path $\tilde{\pi}$ in a random walk with parameter $r/(r + 1)$. Thus, as $x \to \infty$,

$$\mathbb{P}_{\mathrm{CA},r}^{(x,x)}[\tau_n = 2k] = \sum_{\tilde{\pi} \in A_{n,2k}} \mathbb{P}_{\mathrm{CA},r}^{(x,x)}[x + \tilde{\pi}] \to \sum_{\tilde{\pi} \in A_{n,2k}} \mathbb{P}_{\mathrm{RW},r}^{(0,0)}[\tilde{\pi}] = \mathbb{P}_{\mathrm{RW},r}^{(0,0)}[\tau_n = 2k].$$

After summing over k and using Fatou's Lemma, we obtain

$$\mathbb{P}_{\text{CA},r}^{(x,x)}[N \geq n] = \mathbb{P}_{\text{CA},r}^{(x,x)}[\tau_n < \infty] = \sum_{k=1}^{\infty} \mathbb{P}_{\text{CA},r}^{(x,x)}[\tau_n = 2k] \gtrsim \sum_{k=1}^{\infty} \mathbb{P}_{\text{RW},r}^{(0,0)}[\tau_n = 2k] = \mathbb{P}_{\text{RW},r}^{(0,0)}[\tau_n < \infty] = \mathbb{P}_{\text{RW},r}^{(0,0)}[N \geq n].$$

Since $\mathbb{P}_{\text{CA},r}^{(x,x)}[N \geq n] \leq \mathbb{P}_{\text{RW},r}^{(0,0)}[N \geq n]$ by Theorem 4 (whose proof does not rely on the present lemma), we obtain

$$\mathbb{P}_{\text{CA},r}^{(x,x)}[N \geq n] \to \mathbb{P}_{\text{RW},r}^{(0,0)}[N \geq n].$$

In particular,

$$q_r(x,x) = 1 - \mathbb{P}_{\text{CA},r}^{(x,x)}[N \geq 2] \to 1 - \mathbb{P}_{\text{RW},r}^{(0,0)}[N \geq 2] = \frac{r-1}{r+1},$$

completing the proof. $\qquad\square$

3.   Proof of Theorem 3

Recall that $A_{n,t}(x_0, y_0)$ is the set of paths starting from (x_0, y_0) that end with the n-th tie at time t, i.e., $\tau_n = t$. We will use the short-hand notation $A_{n,t}$ for $A_{n,t}(x_0, y_0)$. As in Section A 1, the set $A_{n,t}$ is non-empty only if $t = |x_0 - y_0| + 2k$ for some integer $k \geq n-1$, in which case, every path in $A_{n,t}$ ends in state $(z_0 + k, z_0 + k)$ with $z_0 = \max\{x_0, y_0\}$. Recall from [25] that the probability of any path π connecting states (x_0, y_0) and (x, y) is

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[\pi] = \frac{B(x,y)}{B(x_0,y_0)} = \frac{B(x,y)}{B(x_0,y_0)} 2^t \mathbb{P}_{\text{RW},1}^{(x_0,y_0)}[\pi].$$

Summing over $\pi \in A_{n,t}$, where $x = y = z_0 + k$, we obtain

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[\tau_n = t] = \frac{B(z_0 + k, z_0 + k)}{B(x_0, y_0)} 2^t \mathbb{P}_{\text{RW},1}^{(x_0,y_0)}[\tau_n = t]. \tag{A11}$$

Thus the probability of having the n-th and also the last tie at time t is given by

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[T = t, N = n] = \mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[\tau_n = t] \cdot q_1(z_0 + k, z_0 + k) = \frac{1}{2^{x_0+y_0-2}B(x_0, y_0)} \cdot \frac{1}{t + x_0 + y_0} \mathbb{P}_{\text{RW},1}^{(x_0,y_0)}[\tau_n = t] \tag{A12}$$

where we have used Eqs. (A11) and (A3) in the last step. Note that $\mathbb{P}_{\text{RW},1}^{(x_0,y_0)}[\tau_n = t]$ is the probability $f_{n,t}(d_0)$ of the n-th visit to the origin at time t in a simple symmetric random walk starting from $d_0 = |x_0 - y_0|$. Summing over t in (A12), we obtain

$$\mathbb{P}_{\text{CA},1}^{(x_0,y_0)}[N = n] = \frac{1}{2^{x_0+y_0-2}B(x_0, y_0)} \sum_{k=n-1}^{\infty} \frac{f_{n,d_0+2k}(d_0)}{2k + d_0 + x_0 + y_0} = \frac{G_n(1; d_0)}{2^{x_0+y_0-2}B(x_0, y_0)}, \tag{A13}$$

where

$$G_n(z; d_0) = \sum_{k=n-1}^{\infty} \frac{f_{n,d_0+2k}(d_0)}{2k + d_0 + x_0 + y_0} z^{d_0+2k}.$$

To simplify $G_n(z; d_0)$, we have

$$\frac{d}{dz}[z^{x_0+y_0} G_n(z; d_0)] = z^{x_0+y_0-1} \sum_{k=n-1}^{\infty} f_{n,d_0+2k}(d_0) z^{d_0+2k} = z^{x_0+y_0-1} \Phi_n(z; d_0), \tag{A14}$$

where $\Phi_n(z; d_0) = \sum_{k=n-1}^{\infty} f_{n,d_0+2k}(d_0) z^{d_0+2k}$ is the generating function of the probability distribution of the n-th visit to the origin in a simple random walk starting from d_0. Let $F_1(z)$ be the generating function of the distribution of the time of the first return to the origin in a simple random walk starting from the origin. The standard renewal argument (see e.g. XI.3.d of [35]) shows that $\Phi_n(z; d_0)$ is given by

$$\Phi_n(z; d_0) = [\Phi_1(z; 1)]^{d_0} [F_1(z)]^{n-1},$$

where $\Phi_1(z;1)$ and $F_1(z)$ are given by Eqs. (3.6) and (3.14) of [35, Chap. XI], respectively. Therefore,

$$\Phi_n(z;d_0) = z^{-d_0}\left(1 - \sqrt{1 - z^2}\right)^{n+d_0-1}. \tag{A15}$$

Substituting Eq. (A15) into Eq. (A14) and integrating from 0 to 1 yields

$$G_n(1;d_0) = \int_0^1 z^{2\min\{x_0,y_0\}-1}\left(1 - \sqrt{1 - z^2}\right)^{n+d_0-1}dz,$$

where we have used $x_0 + y_0 - d_0 = 2\min\{x_0, y_0\}$. A change of variable $u = \sqrt{1 - z^2}$ yields

$$G_n(1;d_0) = \int_0^1 u(1 - u^2)^{\min\{x_0,y_0\}-1}(1 - u)^{n+d_0-1}du,$$

which is upper bounded by

$$G_n(1;d_0) \leq \int_0^1 u(1 - u)^{n+d_0-1}du = B(2, n + d_0), \tag{A16}$$

and lower bounded by

$$G_n(1;d_0) \geq \int_0^1 u(1 - u)^{\min\{x_0,y_0\}-1}(1 - u)^{n+d_0-1}du = B\left(2, n + \max\{x_0, y_0\} - 1\right), \tag{A17}$$

where we have used $\min\{x_0, y_0\} + d_0 = \max\{x_0, y_0\}$. Applying Eqs. (A16) and (A17) to Eq. (A13) yields

$$\frac{B(2, n + \max\{x_0, y_0\} - 1)}{2^{x_0+y_0-2}B(x_0, y_0)} \leq \mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[N = n] \leq \frac{B(2, n + d_0)}{2^{x_0+y_0-2}B(x_0, y_0)}.$$

Note that $\mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[N = \infty] = \mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[T = \infty] = 0$. Summing over n and using $\sum_{m=n}^{\infty} B(2, m) = n^{-1}$, we obtain

$$\frac{1}{2^{x_0+y_0-2}B(x_0, y_0)} \cdot \frac{1}{n + \max\{x_0, y_0\} - 1} \leq \mathbb{P}_{\mathrm{CA},1}^{(x_0,y_0)}[N \geq n] \leq \frac{1}{2^{x_0+y_0-2}B(x_0, y_0)} \cdot \frac{1}{n + d_0},$$

which immediately yields Eq. (6).

4. Proof of Theorem 4

We first prove the following lemma.

Lemma 4. *The probability* $\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty]$ *of ever having a tie is bounded as follows,*

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty] \leq \begin{cases} 1, & x_0 \leq y_0, \\ \frac{(y_0)_{x_0-y_0}}{(rx_0+y_0)_{x_0-y_0}}\left(1 + \frac{1}{r}\right)^{x_0-y_0}, & x_0 > y_0. \end{cases}$$

Note that $\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty]$ is the exit probability $E(x, y)$ in [25] when $r = 1$.

Proof. The case $x_0 = y_0$ is trivial since $\tau_1 = 0$. When $x_0 < y_0$, Theorem 3.3 of [21] yields $Y_t/X_t \to 0$ almost surely, from which it follows that $X_t > Y_t$ eventually and hence $\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty] = 1$.

Now assume $x_0 > y_0$. Recall that $A_{1,t}(x_0, y_0)$ is the set of paths starting from (x_0, y_0) that end with the first tie at time t. Note that $A_{1,t}(x_0, y_0)$ is nonempty only if $t = d_0 + 2k$, where $d_0 = x_0 - y_0$ and $k \geq 0$. Let $\pi \in A_{1,t}(x_0, y_0)$ and its state at time j be $\pi_j = (x_j, y_j)$. The probability of the path π is given by

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\pi] = \prod_{j=0}^{t-1}\left(\frac{rx_j}{rx_j + y_j}\right)^{x_{j+1}-x_j}\left(\frac{y_j}{rx_j + y_j}\right)^{y_{j+1}-y_j} = \frac{r^{x_t-x_0}(x_0)_{x_t-x_0}(y_0)_{y_t-y_0}}{\prod_{j=0}^{t-1}(rx_j + y_j)} = \frac{r^{x_t-x_0}(x_0)_{x_t-x_0}(y_0)_{y_t-y_0}}{\prod_{j=0}^{t-1}[(r - 1)x_j + x_0 + y_0 + j]},$$

where in the last step we have used $x_j + y_j = x_0 + y_0 + j$. Note that $\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\pi]$ is maximized if the x_j's are minimized, subject to the constraints that the x_j's increase monotonically from x_0 to x_t with step size 0 or 1, and that $x_j > y_j$ for all $1 \le j \le t-1$, or equivalently $x_j > x_0 + (j - d_0)/2$. This is achieved by the following sequence,

$$x_j^* = \begin{cases} x_0, & j = 0, 1, \ldots, d_0 - 1; \\ x_0 + \lfloor (j - d_0)/2 \rfloor + 1, & j = d_0, d_0 + 1, \ldots, t - 1; \\ x_t, & j = t. \end{cases}$$

The corresponding path π^* has probability

$$\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\pi^*] = \prod_{j=0}^{d_0-2} \frac{y_0 + j}{rx_0 + y_0 + j} \prod_{x=x_0}^{x_t-1} \frac{rx}{rx + (x-1)} \cdot \frac{x-1}{r(x+1) + (x-1)} \frac{x_t - 1}{rx_t + (x_t - 1)},$$

which, after arrangement, yields,

$$\begin{aligned} \mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\pi^*] &= \prod_{j=0}^{d_0-1} \frac{y_0 + j}{rx_0 + y_0 + j} \prod_{x=x_0}^{x_t-1} \frac{rx}{(r+1)x + r} \cdot \frac{x}{(r+1)x + (r-1)} \\ &\le \frac{(y_0)_{d_0}}{(rx_0 + y_0)_{d_0}} \frac{r^{x_t - x_0}}{(r+1)^{2(x_t - x_0)}} \\ &= \frac{(y_0)_{d_0}(r+1)^{d_0}}{(rx_0 + y_0)_{d_0}} \left(\frac{r}{r+1}\right)^{x_t - x_0} \left(\frac{1}{r+1}\right)^{y_t - y_0} \\ &= \frac{(y_0)_{d_0}(r+1)^{d_0}}{(rx_0 + y_0)_{d_0}} \mathbb{P}_{\text{RW},r}^{(x_0,y_0)}[\pi]. \end{aligned}$$

Thus we have

$$\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\pi] \le \mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\pi^*] \le \frac{(y_0)_{d_0}(r+1)^{d_0}}{(rx_0 + y_0)_{d_0}} \mathbb{P}_{\text{RW},r}^{(x_0,y_0)}[\pi],$$

and, after summing over $\pi \in A_{1,t}(x_0, y_0)$,

$$\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\tau_1 = t] \le \frac{(y_0)_{d_0}(r+1)^{d_0}}{(rx_0 + y_0)_{d_0}} \mathbb{P}_{\text{RW},r}^{(x_0,y_0)}[\tau_1 = t].$$

Summing over t, we obtain

$$\mathbb{P}_{\text{CA},r}^{(x_0,y_0)}[\tau_1 < \infty] \le \frac{(y_0)_{d_0}(r+1)^{d_0}}{(rx_0 + y_0)_{d_0}} \mathbb{P}_{\text{RW},r}^{(x_0,y_0)}[\tau_1 < \infty].$$

By Eq. (XI.3.9) and XI.3.d in [35], $\mathbb{P}_{\text{RW},r}^{(x_0,y_0)}[\tau_1 < \infty] = r^{-d_0}$, from which the desired conclusion follows. $\square$

Corollary 1. *The probability of having at least one more tie starting from a tie state* (x, x) *is bounded by*

$$\mathbb{P}_{\text{CA},r}^{(x,x)}[\tau_2 < \infty] \le \frac{2}{r+1}.$$

Proof. By considering the one-step transition from (x, x) into $(x, x + 1)$ or $(x + 1, x)$, we obtain

$$\mathbb{P}_{\text{CA},r}^{(x,x)}[\tau_2 < \infty] = \frac{r}{r+1} \mathbb{P}_{\text{CA},r}^{(x+1,x)}[\tau_1 < \infty] + \frac{1}{r+1} \mathbb{P}_{\text{CA},r}^{(x,x+1)}[\tau_1 < \infty] \le \frac{x}{(r+1)x + r} + \frac{1}{r+1} \le \frac{2}{r+1},$$

where the first inequality follows from Lemma 4. $\square$

Now we prove Theorem 4.

Proof of Theorem 4. Let Z_n be the common value of X_t and Y_t at $t = \tau_n$, i.e., $Z_n = X_{\tau_n}$. Conditioned on $\tau_n < \infty$ and $Z_n = z$, the probability of $\tau_{n+1} < \infty$ is just the probability of having a tie after leaving (z, z). Thus

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_{n+1} < \infty \mid \tau_n < \infty, Z_n = z] = \mathbb{P}_{\mathrm{CA},r}^{(z,z)}[\tau_2 < \infty] \leq \frac{2}{r+1},$$

by Corollary 1. Removal of the conditioning yields

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_{n+1} < \infty \mid \tau_n < \infty] \leq \frac{2}{r+1}.$$

It follows that

$$\mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[N \geq n] = \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_n < \infty] = \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty] \prod_{i=1}^{n-1} \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_{i+1} < \infty \mid \tau_i < \infty] \leq \mathbb{P}_{\mathrm{CA},r}^{(x_0,y_0)}[\tau_1 < \infty] \left(\frac{2}{r+1}\right)^{n-1}.$$

An application of Lemma 4 completes the proof. $\qquad\square$