Balls-in-bins with feedback and Brownian Motion

Roberto Oliveira*

January 8, 2007

Abstract

In a balls-in-bins process with feedback, balls are sequentially thrown into bins so that the probability that a bin with n balls obtains the next ball is proportional to $f(n)$ for some function f. A commonly studied case where there are two bins and $f(n) = n^p$ for $p > 0$, and our goal is to study the fine behavior of this process with two bins and a large initial number t of balls. Perhaps surprisingly, Brownian Motions are an essential part of both our proofs.

For $p > 1/2$, it was known that with probability 1 one of the bins will lead the process at all large enough times. We show that if the first bin starts with $t + \lambda\sqrt{t}$ balls (for constant $\lambda \in \mathbf{R}$), the probability that it always or eventually leads has a non-trivial limit depending on λ.

For $p \leq 1/2$, it was known that with probability 1 the bins will alternate in leadership. We show, however, that if the initial fraction of balls in one of the bins is $> 1/2$, the time until it is overtaken by the remaining bin scales like $\Theta(t^{1+1/(1-2p)})$ for $p < 1/2$ and $\exp(\Theta(t))$ for $p = 1/2$. In fact, the overtaking time has a non-trivial distribution around the scaling factors, which we determine explicitly.

Our proofs use a continuous-time embedding of the balls-in-bins process (due to Rubin) and a non-standard approximation of the process by Brownian Motion. The techniques presented also extend to more general functions f.

1 Introduction

1.1 The process

If $f : \mathbf{N} \to (0, +\infty)$ is a positive function, a *balls-in-bins process with feedback function* f a discrete-time Markov process with B bins, each one of which containing $I_i(m) > 0$ balls at time m for each $m \in \{0, 1, 2, \ldots\}$ and $i \in \{1, \ldots, B\}$. Its evolution is as follows: at each time

*IBM T.J. Watson Research Center, Yorktown Heights, NY 10598. `rioliveі@us.ibm.com`. Work mostly done while the author was a Ph.D. student at New York University under a CNPq scholarship.

$m > 0$, a ball is added to a bin i_m, so that $I_{i_m}(m) = I_{i_m}(m-1)+1$ and $I_j(m) = I_j(m-1)$ for all $i \in \{1,\dots,B\}\backslash\{i_m\}$, and the random choice of bin i_m has distribution

$$\mathbf{Pr}\left(i_m = i \mid \{I_j(m-1)\,:\,1\leq j\leq B\}\right) = \frac{f(I_i(m-1))}{\sum_{j=1}^{B} f(I_j(m-1))}\ \ (1\leq i\leq B), \tag{1}$$

We will commonly refer to this recipe by saying that *bin i receives a ball at time m (i.e. $i_m = i$) with probability proportional to $f(I_m(i))$.*

Such processes[1] were introduced to the Discrete Mathematics community by Drinea, Frieze and Mitzenmacher [6], where they were motivated by economical problems of competition and mathematically related preferential attachment models for large networks[2]. That paper treats only the case where $f(x) = x^p$ for some parameter $p > 0$. In this case f is increasing, and therefore *the rich get richer*: the more balls a bin has, the more likely it is to receive the next ball. In economic terms, one could think of bins as competing products and balls as customers; in that case, the more popular a product is, the more likely it is to obtain a new customer.

The main question addressed in that paper is whether this phenomenon ensures large advantages *in the long run* for some bin. The authors show that if $p > 1$, there almost surely exists one bin that gets all but a negligible fraction of the balls in the large-time limit; whereas for $p < 1$, the asymptotic fractions of balls in each bin even out with time. The $p = 1$ case is the classic Pólya Urn model, for which it has been long known that the number of balls in each bin converges almost surely to a non-degenerate random variable, and thus the process has different *regimes* depending on the choice of parameter p.

1.2 The three regimes

However, a much stronger result is available. A paper by Khanin and Khanin [7] introduced what amounts to the same process as a model for neuron growth, and proved that if $p > 1$, there almost surely is some bin that gets *all but finitely many balls*, an event that we call *monopoly*. They also show that for $1/2 < p \leq 1$, monopoly has probability 0, but there almost surely will be some bin which will lead the process from some finite time on (we call this *eventual leadership*), whereas this cannot happen if $0 < p \leq 1/2$. Therefore, the balls-in-bins process has *three regimes* of behavior corresponding to three ranges of p.

In fact, the result of [7] generalizes to any f with $\min_{x\in\mathbf{N}} f(x) > 0$. Consider the following events, which we call *monopoly by bin $i \in [B]$*,

$$\begin{aligned}\mathsf{Mon}_i &\equiv \{\exists M\in\mathbf{N}\ \forall m\geq M\ \ \forall j\in[B]\ \ j\neq i \Rightarrow I_m(j) = I_M(j)\} \\ &= \{\exists M\in\mathbf{N}\ \forall m\geq M\ \ i_m = i\};\end{aligned} \tag{2}$$

[1] A longer background discussion is available from the author's PhD thesis [10].

[2] More specifically, the model introduced by Krapivsky and Redner [8] and independently by Driena, Enachescu and Mitzemacher [5]. This model generalizes the Barabási-Álbert model, which is discussed in the survey [1].

and *eventual leadership by bin* $i \in [B]$:

$$\mathsf{ELead}_i \equiv \{\exists M \in \mathbf{N}\ \forall m \geq M\ \ \forall j \in [B]\ \ j \neq i \Rightarrow I_m(j) < I_m(i)\}. \tag{3}$$

Clearly, $\mathsf{ELead}_i \supset \mathsf{Mon}_i$. A not-too-difficult extension of [7] (proven in [13, 10]) says that

Theorem 1 (From [7, 13, 10]) *If $\{I_m\}_{m=0}^{+\infty}$ is a balls-in-bins process with B bins and feedback function $f = f(x) \geq c$ for some $c > 0$, then there are three mutually exclusive possibilities.*

1. $\sum_{n\geq 1} f(n)^{-1} < +\infty$*, in which case* $\mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{Mon}_i\right) = \mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{ELead}_i\right) = 1$ *(we call this the* monopolistic *regime);*
2. $\sum_{n\geq 1} f(n)^{-1} = +\infty$ *but* $\sum_{n\geq 1} f(n)^{-2} < +\infty$*, in which case* $\mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{Mon}_i\right) = 0$ *but* $\mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{ELead}_i\right) = 1$ *(this is the* eventual leadership regime*);*
3. $\sum_{n\geq 1} f(n)^{-2} = +\infty$*, in which case* $\mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{Mon}_i\right) = \mathbf{Pr}\left(\cup_{i=1}^{B} \mathsf{ELead}_i\right) = 0$ *(this is the* almost-balanced regime*).*

This holds irrespective of the initial conditions of the process.

Notice that the three cases of the Theorem applied to the $f(x) = x^p$ family correspond to $p > 1$, $1/2 < p \leq 1$ and $0 \leq p < 1/2$; in other words, this family of f has *phase transitions* at $p = 1$ and $p = 1/2$. We sketch a proof of this result in Section 3.2, both for completeness and to give readers a better acquaintance with the techniques in the present paper.

1.3 The present work

This paper is part of a series by the present author in collaboration with Michael Mitzenmacher and Joel Spencer [9, 11, 12] that attempts to quantify different aspects of the three qualitative regimes presented in Theorem 1. Our specific purpose in the present paper is to prove two not-quite-related results about these processes in different regimes, when the initial number of balls on both bins is large. What brings these two results is that both proofs use Brownian Motion in an unexpected and surprising way.

Our first result is a *scaling result for the eventual leadership and monopoly regimes.* Suppose, for simplicity, that $f(x) = x^p$ with $p > 1/2$. Recall the definition of eventual leadership by bin i (3), and let Lead_i be the event that bin 1 leads the process at *all times*:

$$\mathsf{Lead}_i \equiv \{\forall m \geq 0\ \ \forall j \in [B]\ \ j \neq i \Rightarrow I_m(j) < I_m(i)\}. \tag{4}$$

If the initial number of bins is $t \gg 1$ and $I_0(1) \approx t/2$, then $\mathbf{Pr}\left(\mathsf{Lead}_i\right) \approx 0$ and $\mathbf{Pr}\left(\mathsf{ELead}_i\right) \approx 1$ for $i = 1, 2$. On the other hand, if $I_0(1)$ is much larger than $I_0(2)$, one think that $\mathbf{Pr}\left(\mathsf{Lead}_1\right), \mathbf{Pr}\left(\mathsf{ELead}_1\right) \approx 1$. Our question is *how large is large enough*? That is to say, at what *scale* do these two probabilities grow from 0 to 1? We show that the answer is in fact $\Theta\left(\sqrt{t}\right)$, and give an exact asymptotic result.

Theorem 2 *Let* $\{\lambda_t\}_{t\in\mathbf{N}} \subset \mathbf{R}$ *form a sequence such that such that*

$$\lambda \equiv \lim_{t\to+\infty} \lambda_t \in \mathbf{R} \textit{ exists}$$

and

$$\forall t \in \mathbf{N},\ \ \frac{t}{2} \pm \lambda_t \sqrt{\frac{t}{4p-2}} \in \mathbf{N}.$$

Assume that for each t, $\{I_m^{(t)}(j)\ :\ j \in [2], m \geq 0\}$ *is a balls-in-bins process with two bins, feedback function* $f(x) = x^p$, $1/2 < p < +\infty$ *and initial conditions*

$$(x(t), y(t)) \equiv \left(\frac{t}{2} + \lambda_t\sqrt{\frac{t}{4p-2}}, \frac{t}{2} - \lambda_t\sqrt{\frac{t}{4p-2}}\right).$$

Then

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x(t),y(t))}(\mathsf{ELead}_1) = \gamma(\lambda) \equiv \sqrt{\frac{1}{2\pi}}\int_{-\infty}^{\lambda} e^{-x^2/2}\,dx, \quad (5)$$

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x(t),y(t))}(\mathsf{Lead}_1) = \Gamma(\lambda) \equiv \sqrt{\frac{2}{\pi}}\int_{0}^{\max\{\lambda,0\}} e^{-x^2/2}\,dx. \quad (6)$$

This theorem is an extension of a result by Mitzenmacher, Oliveira and Spencer [9], who showed a similar scaling for Mon_1 when $p > 1$. That paper used Esséen's inequality for approximation by Gaussians together with a continuous-time embedding of the balls-in-bins process; we shall also use the latter device together with approximation by Brownian Motion, especially to estimate $\mathbf{Pr}_{(x(t),y(t))}(\mathsf{Mon}_i)$. Notice that as $p \searrow 1/2$ the scaling term $\sqrt{\frac{t}{4p-2}}$ becomes bigger; i.e. near the $p = 1/2$ phase transition, it becomes harder to bias the process towards (eventual) leadership by either bin.

The second result we prove is about the almost-balanced case. Suppose, again for simplicity, that there are two bins ($B = 2$) and $f(x) = x^p$, $0 < p \leq 1/2$, . In this case Theorem 1 says that for any initial conditions $I_0(1), I_0(2)$ with $I_0(1) < I_0(2)$, there is a time $m \geq 0$ such that $I_m(1) > I_m(2)$. Call the first such time the *overtaking time* V. By the above, $V < +\infty$, but we have no idea of the distribution of V, and thus we don't know how long the overtaking might take. We show that if the initial number of balls is large and bin 2 has a non-negligibly bigger fraction of the initial balls, then V can actually be quite large; moreover, it has an explicit asymptotic distribution.

Theorem 3 *Let* $V_{t,\alpha}$ *be the overtaking time in a balls-in-bins process with feedback function* $f(x) = x^p$ *(with* $p \in (0, 1/2]$ *constant) and initial conditions* $(\lceil \alpha t\rceil, t - \lceil \alpha t\rceil)$ *for* $0 < \alpha < 1/2$.

Then there exist random variables $\{U_{t,\alpha,p}\}_{t\in\mathbf{N}}$ *such that*

$$V_{t,\alpha} = \begin{cases} 2\left\{\frac{(1-2p)}{(1-p)^2}[(1-\alpha)^{1-p}-\alpha^{1-p}]^2 + \frac{(1-\alpha)^{1-2p}\,U_{t,\alpha,p}^2}{t}\right\}^{\frac{1}{1-2p}} \frac{t^{1+\frac{1}{1-2p}}}{U_{t,\alpha,p}^{1+\frac{2p}{1-2p}}} - (t+1) \\ \qquad \textit{if } 0 < p < \frac{1}{2}, \\ 2(1-\alpha)\,t\,\exp\left\{4[1-2\sqrt{\alpha(1-\alpha)}]\,\frac{t}{U_{t,\alpha,\frac{1}{2}}^2}\right\} - (t+1) \\ \qquad \textit{if } p = \frac{1}{2}, \end{cases} \tag{7}$$

with probability tending to 1 *as* $t \to +\infty$, *and*

$$U_{t,\alpha,p} \to^{w} |N|,$$

where N *is a standard Gaussian random variable.*

This means that V becomes larger and larger as $p \nearrow 1/2$, culminating with the exponential behavior at the phase transition point $p = 1/2$. The economically-inclined might wish to deduce from this theorem that, under appropriate initial conditions, a product's leadership might last a long time even in markets with no propensity for breeding monopolies or "eternal leaders".

1.4 Techniques and outline

Our results in this paper are actually more general: they extend to a broader (though not entirely general) class of functions f and, in the case of Theorem 6, to more than two bins. All proofs below are done for this more general case and then specialized for $f(x) = x^p$.

Our proofs have their first two steps in common. The first step has been employed in [7, 13] and other works, and seems to have originated in Davis' work on reinforced random walks [4]. We shall *embed* the discrete-time process we are interested in into a continuous-time process built from exponentially distributed random variables, so that interarrival times at different bins are independent and have an explicit distribution, which is very helpful in calculations. We call this the *exponential embedding* of the process.

The second technique we use is *approximation by Brownian motion* via Donsker's Invariance Principle. While neither technique is novel, their conjunction in the way presented here yields surprising explicit results in the asymptotic regime, once the appropriate calculations are done.

The remainder of the paper is organized as follows. We discuss preliminary material in Section 2. Section 3 rigorously introduces the exponential embedding process and discusses its key properties. In Section 4 we detail the assumptions we make on our feedback functions f, while also deriving some consequences of those assumptions. Section 6 proves the general version of Theorem 2, whereas Section 7 contains the proof of the generalization of Theorem 3. Finally, Section 8 discusses extensions to our results and some open problems.

1.5 Acknowledgements

Michael Mitzenmacher and my Ph.D. advisor Joel Spencer introduced me to this topic and stimulated me with several interesting questions and feedback. I also thank Eleni Drinea for useful discussions.

2 Preliminaries

General notation. Throughout the paper, $\mathbf{N} = \{1, 2, 3, \ldots\}$ is the set of non-negative integers, $\mathbf{R}^+ = [0, +\infty)$ is the set of non-negative reals, and for any $k \in \mathbf{N}\backslash\{0\}$ $[k] = \{1, \ldots, k\}$.

Asymptotics. We use the standard $O/o/\Omega/\Theta$ notation. The expressions "$a_n \sim b_n$ as $n \to n_0$" and "$a_n \ll b_n$ as $n \to n_0$" mean that $\lim_{n\to n_0}(a_n/b_n) = 1$ and $\lim_{n\to n_0}(a_n/b_n) = 0$, respectively.

Balls-in-bins. Formally, a balls-in-bins process with feedback function $f : \mathbf{N} \to (0 + \infty)$ and $B \in \mathbf{N}$ bins is a discrete-time Markov chain $\{(I_1(m), \ldots, I_B(m))\}_{m=0}^{+\infty}$ with state space $\mathbf{N}^B$ and transition probabilities given in the Introduction (see (1)). We will usually refer to the index $i_m \in [B]$ as *the bin that receives a ball at time m*. For any B, if E is an event of the process and $u \in \mathbf{N}^B$, $\mathbf{Pr}_u(E)$ is the probability of E when the initial conditions are set to u.

Exponential random variables. $X =^d \exp(\lambda)$ means that X is a random variable with exponential distribution with rate $\lambda > 0$, meaning that $X \geq 0$ and

$$\mathbf{Pr}(X > t) = e^{-\lambda t} \quad (t \geq 0).$$

The shorthand $\exp(\lambda)$ will also denote a generic random variable with that distribution. Some elementary but extremely useful properties of those random variables include

1. *Lack of memory.* Let $X =^d \exp(\lambda)$ and $Z \geq 0$ be independent from X. The distribution of $X - Z$ conditioned on $X > Z$ is still equal to $\exp(\lambda)$.

2. *Minimum property.* Let $\{X_i =^d \exp(\lambda_i)\}_{i=1}^m$ be independent. Then

   $$X_{\min} \equiv \min_{1\leq i\leq m} X_i =^d \exp(\lambda_1 + \lambda_2 + \ldots \lambda_m)$$

   and for all $1 \leq i \leq m$

   $$\mathbf{Pr}(X_i = X_{\min}) = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \ldots \lambda_m} \tag{8}$$

3. *Multiplication property.* If $X =^d \exp(\lambda)$ and $\eta > 0$ is a fixed number, $\eta X =^d \exp(\lambda/\eta)$.

4. *Moments and transforms.* If $X =^d \exp(\lambda)$, $r \in \mathbf{N}$ and $t \in \mathbf{R}$,

$$\begin{aligned} \mathbf{Ex}[X^r] &= \frac{r!}{\lambda^r}, \qquad (9) \\ \mathbf{Ex}\left[e^{tX}\right] &= \begin{cases} \frac{1}{1-\frac{t}{\lambda}} & (t < \lambda) \\ +\infty & (t \geq \lambda) \end{cases} \qquad (10) \end{aligned}$$

Weak convergence. $X_n \to^w Y$ means that the sequence $\{X_n\}$ of random variables converges weakly to Y as $n \to +\infty$.

Gaussians and cumulative distribution functions. Finally, we restate the definitions of Φ and Γ in Theorem 6:

$$\begin{aligned} \Phi(\lambda) &\equiv \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\lambda} e^{-x^2/2}\, dx, \qquad (11) \\ \Gamma(\lambda) &\equiv \sqrt{\frac{2}{\pi}} \int_{0}^{\max\{\lambda,0\}} e^{-x^2/2}\, dx. \qquad (12) \end{aligned}$$

If N is a standard Gaussian random variable, Γ is the cdf (cumulative distribution function) of N and Γ is the cdf of $|N|$.

3 The exponential embedding

3.1 Definition and key properties

Let $f : \mathbf{N} \to (0, +\infty)$ be a function, $B \in \mathbf{N}$ and $(a_1, \dots, a_B) \in \mathbf{N}^B$. We define below a continuous-time process with state space $(\mathbf{N} \cup \{+\infty\})^B$ and initial state $(a_1, \dots, a_B)$ as follows. Consider a set $\{X(i,j) : i \in [B], j \in \mathbf{N}\}$ of independent random variables, with $X(i,j) =^d \exp(f(j))$ for all $(i,j) \in [B] \times \mathbf{N}$, and define

$$N_i(t) \equiv \sup\left\{ n \in \mathbf{N} : \sum_{j=a_i}^{n-1} X(i,j) \leq t \right\} \quad (i \in [B], t \in \mathbf{R}^+ = [0, +\infty)), \qquad (13)$$

where by definition $\sum_{j=i}^{k}(\dots) = 0$ if $i > k$. Thus $N_i(0) = a_i$ for each $i \in [B]$, and one could well have $N_i(T) = +\infty$ for some finite time T (indeed, that *will* happen for our cases of interest); but in any case, the above defines a continuous-time stochastic process, and in fact the $\{N_i(\cdot)\}_{i=1}^{B}$ processes are independent. Each one of this processes is said to correspond to *bin* i, and each one of the times

$$X(i, a_i), X(i, a_i) + X(i, a_i + 1), X(i, a_i) + X(i, a_i + 1) + X(i, a_i + 2), \dots$$

is said to be an *arrival time at bin i*. As in the balls-in-bins process, we imagine that each arrival correspond to a ball being placed in bin i.

In fact, we *claim* that this process is related as follows to the balls-in-bins process with feedback function f, B bins and initial conditions $(a_1, \dots, a_B)$.

Theorem 4 (Proven in [4, 7, 13, 10, 11]) *Let the $\{N_i(\cdot)\}_{i\in[B]}$ process be defined as above. One can order the arrival times of the B bins in increasing order (up to their first accumulation point, if they do accumulate) so that $T_1 < T_2 < \dots$ is the resulting sequence. The distribution of*

$$\{I_m = (N_1(T_m), N_2(T_m), \dots, N_B(T_m))\}_{m\in\mathbf{N}}$$

is the same as that of a balls-in-bins process with feedback function f and initial conditions $(a_1, a_2, \dots, a_B)$.

One can prove this result[3] as follows. First, notice that the *first arrival time* T_1 is the minimum of $X(j, a_j)$, $(1 \le j \le B)$. By the minimum property presented above, the probability that bin i is the one at which the arrival happens is like the first arrival probability in the corresponding balls-in-bins process with feedback:

$$\mathbf{Pr}\left(X(i, a_i) = \min_{1\le j\le B} X(j, a_j)\right) = \frac{f(a_i)}{\sum_{j=1}^{B} f(a_j)}. \tag{14}$$

More generally, let $t \in \mathbf{R}^+$ and condition on $(N_i(t))_{i=1}^B = (b_i)_{i=1}^B \in \mathbf{N}^B$, with $b_i \ge a_i$ for each i (in which case the process has not blown up). This amounts to conditioning on

$$\forall i \in [B] \sum_{j=a_i}^{b_i-1} X(i, b_i) \le t < \sum_{j=a_i}^{b_i} X(i, b_i).$$

From the lack of memory property of exponentials, one can deduce that the first arrival after time t at a given bin i will happen at a $\exp(f(b_j))$-distributed time, independently for different bins. This almost takes us back to the situation of (14), with b_i replacing a_i, and we can similarly deduce that bin i gets the next ball with the desired probability,

$$\frac{f(b_i)}{\sum_{j=1}^{B} f(b_j)}.$$

[3]The exact attribution of this result is somewhat confusing. Ref. [7] cites the work of Davis [4] on reinforced random walks, where it is in turn attributed to Rubin.

3.2 On the three regimes

Let us now briefly point out some of the key steps in the proof of Theorem 1 via the exponential embedding, in the case $B = 2$. Reading this sketch might help the reader to become acquainted with an important part of our methods.

We use the same notation and random variables introduced above.

Assume we start the process from state $(x, y) \in \mathbf{N}^2$. First, we note that

$$\mathbf{Ex}\left[\sum_{j=x}^{+\infty} X(1,j)\right] = \sum_{j=x}^{+\infty} \frac{1}{f(j)}, \tag{15}$$

hence if the RHS is finite, $\sum_{j=x}^{+\infty} X(1,j) < +\infty$ almost surely, and similarly for $\sum_{j=y}^{+\infty} X(2,j)$. Moreover, the two random series are independent, and neither has point-masses in their distribution. Therefore, with probability 1,

$$\text{either } \sum_{j=x}^{+\infty} X(1,j) < \sum_{j=y}^{+\infty} X(2,j) \text{ or } \sum_{j=x}^{+\infty} X(1,j) > \sum_{j=y}^{+\infty} X(2,j). \tag{16}$$

If the first alternative holds, there exists a finite $M > y$ such that

$$\sum_{j=x}^{+\infty} X(1,j) < \sum_{j=y}^{M-1} X(2,j).$$

Now notice that the sequence $\{T_{n_k} = \sum_{j=x}^{x+k} X(1,j)\}_{k\in\mathbf{N}}$ is an infinite subsequence of the ball arrival times $\{T_n\}_{n\in\mathbf{N}}$, and at those times

$$\forall k \in \mathbf{N} \;\; T_{n_k} = \sum_{j=x}^{x+k} X(1,j) < \sum_{j=y}^{M-1} X(2,j) \Rightarrow I_{n_k}(2) < M.$$

Since $\{I_m(2)\}_m$ is an increasing sequence, this means that $I_m(2) < M$ for all $m \in \mathbf{N}$; that is to say, bin 1 *must achieve* monopoly. On the other hand, if the second alternative in (16) holds, the same argument shows that bin 2 must achieve monopoly. Thus the condition (15) implies that with probability 1, one of the two bins achieves monopoly. It is not too hard to prove that if (15) does *not* hold, then almost surely $\sum_{j=x}^{+\infty} X(1,j) = \sum_{j=y}^{+\infty} X(2,j) = +\infty$; in fact, it suffices to show that, for some $\rho > 0$

$$\mathbf{Ex}\left[\exp(-\rho \sum_{j=x}^{x+k} X(1,j))\right] = \prod_{j=x}^{x+k} \mathbf{Ex}\left[\exp(-\rho X(1,j))\right] \to 0 \text{ as } k \to +\infty$$

and similarly for $\sum_{j=y}^{+\infty} X(2,j)$. In this case one can show, by reversing the above reasoning, with probability 1 no bin will achieve monopoly.

Now assume that $x > y$ (for simplicity) and

$$\sum_{j=1}^{+\infty} \frac{1}{f(j)^2} < +\infty. \tag{17}$$

In this case, even if $\sum_j f(j)^{-1} = +\infty$, the series

$$\sum_{j=x}^{+\infty} (X(1,j) - X(2,j))$$

is made of independent, centered random variables whose variances satisfy

$$\sum_{j=x}^{+\infty} \mathbf{Var}\left((X(1,j) - X(2,j))\right) = \sum_{j=1}^{+\infty} \frac{2}{f(j)^2} < +\infty.$$

Hence Kolmogorov's Three Series Theorem implies that $\sum_{j=x}^{+\infty}(X(1,j) - X(2,j))$ converges. Following the reasoning developed above, we deduce that almost surely

$$\text{either } \sum_{j=x}^{+\infty}(X(1,j) - X(2,j)) - \sum_{j=y}^{x-1} X(2,j) < 0 \text{ or } \sum_{j=x}^{+\infty}(X(1,j) - X(2,j)) - \sum_{j=y}^{x-1} X(2,j) > 0.$$

In the first case, for all large enough M

$$\sum_{j=x}^{M-1} X(1,j) < \sum_{j=y}^{M-1} X(2,j)$$

and one can check that this means that for all large enough M, bin 1 reaches level M *before* bin 2 does (in the embedded and continuous-time processes): that is, bin 1 achieves eventual leadership. Otherwise, if $\sum_{j=x}^{+\infty}(X(1,j) - X(2,j)) - \sum_{j=y}^{x-1} X(2,j) > 0$, bin 2 is the one that achieves eventual leadership. In either case, what we have discussed up to now proves items 1. and 2. of Theorem 1.

Finally, if

$$\sum_{j=1}^{+\infty} \frac{1}{f(j)^2} = +\infty, \tag{18}$$

then for any x, as $k \to +\infty$,

$$\sum_{j=x}^{x+k} \frac{1}{f(j)^3} \leq \frac{1}{\min_{j \geq 1} f(j)} \sum_{j=x}^{x+k} \frac{1}{f(j)^2} \ll \left(\sum_{j=x}^{x+k} \frac{1}{f(j)^2} \right)^{3/2}.$$

Checking the moments of the $X(i,j)$'s and using the results in Section 5.2 shows that the sums $\sum_{j=x}^{x+k} X(1,j)$, $\sum_{j=y}^{y+k} X(2,j)$ ($k \in \mathbf{N}$) are in the domain of attraction of Brownian Motion for any x and y. This implies that there is a sequence of random numbers $M_1 < M_2 < M_3 < \dots$ and a constant $0 < \alpha < 1$ such that for all $n \in \mathbf{N}$

$$\alpha < \mathbf{Pr}\left(\sum_{j=x}^{k_n} X(1,j) < \sum_{j=y}^{k_n} X(2,j) \,\middle|\, \{k_1, \dots, k_{n-1}\} \cup \{X(1,\ell), X(2,\ell) : \ell \leq k_{n-1}\} \right) < 1-\alpha.$$

This implies that both $\sum_{j=x}^{k_n} X(1,j) < \sum_{j=y}^{k_n} X(2,j)$ and $\sum_{j=x}^{k_n} X(1,j) > \sum_{j=y}^{k_n} X(2,j)$ must occur infinitely often almost surely. In this case, there are infinitely many k for which bin 1 reaches level k before bin 2 does, and vice-versa. It follows that (18) implies that with probability 1 neither bin will achieve eventual leadership, and this proves 3. and the theorem.

Remark 1 *Assume that bin* 1 *achieves monopoly. Then all arrivals of the* continuous-time *process at bin* 2 *after time* $\sum_{j=x}^{+\infty} X(1,j)$ *do not actually happen in the embedded* discrete-time *process* $\{I_m = (I_m(1), I_m(2))\}$. *We call these "ghost events" a* fictitious continuation *of our process. This very useful device is akin to the continuation of a Galton-Watson process beyond its extinction time (see e.g. [2]) and is equally useful in calculations and proofs.*

4 Assumptions on feedback functions

The purpose of this rather technical section is two-fold. First, we spell out the technical assumptions on the feedback function f that we need in our proofs. Nothing seems to actually *require* these assumptions, but they facilitate certain estimates that we employ in the proofs.

Some readers might wish to skip the proofs in this section on a first reading.

4.1 Valid feedback functions

The feedback functions we allow in our results satisfy the following definition.

Definition 1 *A function* $f : \mathbf{N} \to (0, +\infty)$ *with* $f(1) = 1$[4] *is said to be a* valid feedback function *if it can be extended to a piecewise* C^1 *function* $g : \mathbf{R}^+ \cup \{0\} \to (0, +\infty)$ *with the following property: if* $(\ln g(\cdot))'$ *is the right-derivative of* $\ln g$, *and* $h(x) \equiv x(\ln g(x))'$ *(for* $x \in \mathbf{R}^+ \cup \{0\}$*),*

1. $\liminf_{x \to +\infty} h(x) = h_{min} > 0$;

[4] The requirement that $f(1) = 1$ is just a normalization condition, as it does not change the process.

2. $\lim_{x\to+\infty} x^{-1/4}h(x) = 0$;

3. *there exist* $C > 0$ *and* $x_0 \in \mathbf{R}^+$ *such that for all* $\epsilon \in (0,1)$ *and all* $x \geq x_0$

$$\sup_{x\leq t\leq x^{1+\epsilon}} \left|\frac{h(t)}{h(x)} - 1\right| \leq C\,\epsilon. \tag{19}$$

If in addition $h_{min} > 1/2$, *then we say that* f *is ELM (ELM stands for "eventual leadership or monopoly"). If on the other hand* $h(x) \leq 1/2$ *for all large enough* x, *we say* f *is AB ("almost-balanced").*

With slight abuse of notation, we will always assume that f *is defined over* $\mathbf{R}^+ \cup \{0\}$ *and is piecewise* C^1. *We will also call* h *the* characteristic exponent *of* f.

Functions with exponential growth (such as $f(x) = 2^x$) or with oscillations fail to satisfy Definition 1. On the other hand, requiring that f be increasing seems natural, and the smoothness assumption still leaves us with plenty of interesting examples of feedback functions; some examples are given in Table 1 The "canonical case" where $f(x) = x^p$ $(x \geq 1)$ explains the terminology for the characteristic exponent: in that case, $h(x) \equiv p$ for all $x > 1$.

4.2 Consequences of the definition

Let us now define the quantity

$$S_r(n,m) \equiv \sum_{j=n}^{m-1} \frac{1}{f(n)^r} \quad (r \in \mathbf{N}\backslash\{0\};\, n \in \mathbf{N}, m \in \mathbf{N}\cup\{+\infty\}) \tag{20}$$

for some $f : \mathbf{N} \to (0,+\infty)$, and also let $S_r(n) \equiv S_r(n,+\infty)$. If $f(x) = x^p$, then for $m-n, n \gg 1$ a simple shows that

$$S_r(n,m) \sim \int_n^m \frac{dx}{f(x)^r} = \frac{n^{1-rp} - m^{1-rp}}{(rp-1)}.$$

The main content of the following lemmata (the first one proven in [11]) is that a similar result holds for any valid f, if p is replaced by the characteristic exponent h. In particular, any valid f satisfies the monopoly condition in Theorem 1. These lemmas are used in the two main proofs in the paper.

Lemma 1 ([11]) *Assume that* r *is an integer and* f *is a valid feedback function with characteristic exponent* h *satisfying* $h_{\min} > 1/r$. *Define*

$$M_r(n) = \int_n^{+\infty} \frac{dx}{f(x)^r} \quad (r \in \mathbf{N}\backslash\{0\}, n \in \mathbf{N}).$$

Then, as $n \to +\infty$

$$S_r(n) \sim M_r(n) \sim \frac{n}{(rh(n)-1)f(n)^r}.$$

5 Approximation by Brownian Motion

5.1 The Invariance Principle – setup

This is the last section in which technical preliminaries are discussed. In it, we review a form of Donsker's Invariance Principle that shows that under suitable normalization, "nice" partial sums of random variables are close to Brownian Motion. All results in this section are quite standard and can be found in many books on Brownian Motion, e.g. [3]

Consider the vector space $C = C([0,1], \mathbf{R})$ of all real-valued continuous functions on the unit interval, with the sup norm

$$\|\phi(\cdot)\|_{\sup} \equiv \sup_{0\leq s\leq 1} |\phi(s)|, \quad \phi(\cdot) \in C.$$

This gives C a metric and a topology, and from now on we shall think of C as a measurable space with the Borel σ-field. Brownian Motion is simply a probability measure on this measurable space, or rather a random variable $B(\cdot)$ taking values on C, whose defining properties are:

- $\mathbf{Pr}\left(B(0)=0\right)=1$;
- for all $0 \leq s_0 < s_1 < \cdots < s_k \leq 1$, the random variables

  $$\left\{\frac{B(s_i)-B(s_{i-1})}{\sqrt{s_i - s_{i-1}}}\right\}_{i=1}^{k}$$

  are i.i.d. standard Gaussians.

We will also use the following distributional equalities below:

$$\max_{0\leq s\leq 1} B(s),\ -\min_{0\leq t\leq 1} B(t) =^d |N| \text{ where } N \text{ is standard Gaussian;} \tag{21}$$

$$\text{if } B'(\cdot) \text{ is an independent copy of } B,\ \frac{B(\cdot)-B'(\cdot)}{\sqrt{2}} =^d B. \tag{22}$$

Now consider a "triangular sequence" $\{\xi_{n,t}\}_{t\in\mathbf{N}, 1\leq n\leq M_t}$ of independent, 0-mean, square-integrable random variables,. Letting

$$\sigma_{k,t}^2 \equiv \mathbf{Var}\left(\sum_{j=1}^{k} \xi_{j,t}\right) = \sum_{j=1}^{k} \mathbf{Var}\left(\xi_{j,t}\right)\ (k \in [M_t]) \text{ and } \sigma_{0,t}^2 = 0, \tag{23}$$

we define a random element $\Xi_t(\cdot)$ ($t \in \mathbf{N}$) of C as follows.

$$\Xi_t(s) \equiv \frac{\sum_{j=1}^{k(s)} \xi_{j,t} + \left(\frac{s-\sigma^2_{k(t),m}}{\mathbf{Var}(\xi_{k(s)+1,t})}\right)\xi_{k(s)+1,t}}{\sigma_{M_t,t}}, \tag{24}$$

$$\text{where } s \in [0,1] \text{ and} \tag{25}$$

$$k(s) \equiv \max\left\{k \in [M_t] \cup \{0\} : \frac{\sigma^2_{k,t}}{\sigma^2_{M_t,t}} \leq t\right\}. \tag{26}$$

Thus $\Xi_t(\sigma^2_{k,t}/\sigma^2_{M_t,t})$ is the sum of the k first $\xi_{j,t}$'s, divided by a normalizing factor; and for $s \in [\sigma^2_{k,t}/\sigma^2_{M_t,t}, \sigma^2_{k+1,t}/\sigma^2_{M_t,t}]$, $\Xi_t(s)$ is defined by linear interpolation of the values of $\Xi_t(\sigma^2_{k,M_t}/\sigma^2_{t,M_t})$ and $\Xi_t(\sigma^2_{k+1,M_t}/\sigma^2_{t,M_t})$. One can check that this is indeed a measurable element of C.

The Invariance Principle states that if the sequence $\{\xi_{n,M_t}\}$ satisfies certain conditions, the distribution of the $\Xi_t(\cdot)$'s *converges weakly* to a standard Brownian motion $B(\cdot)$. What this means is that if $A \subset C$ is measurable with boundary ∂A and $\mathbf{Pr}\,(B(\cdot) \in \partial A) = 0$, then $\mathbf{Pr}\,(\Xi_t(\cdot) \in A) \to \mathbf{Pr}\,(B(\cdot) \in A)$ as $t \to +\infty$. A sufficient condition for this is given by

Theorem 5 (Special case of Donsker's Invariance Principle) *If*

$$\frac{\sum_{n=1}^{M_t} \mathbf{Ex}\left[|\xi_{n,t}|^3\right]}{\sigma^{3/2}_{M_t,t}} \to 0 \textit{ as } t \to +\infty,$$

then the sequence $\Xi_t(\cdot)$ *converges weakly to* $B(\cdot)$.

5.2 Application to the continuous-time process

Our typical application of this invariance principle will be to the random variables in the exponential embedding. In the notation of the Section 3, let $i, i' \in [B]$ be fixed, $\{x_t\}_t, \{M_t\}_t \subset \mathbf{N}$ be sequences, and consider the triangular array of random variables

$$\left\{\xi_{n,t} \equiv X(i, x_t + n - 1) - X(i', x_t + n - 1)\right\}_{m \in \mathbf{N}, 1 \leq n \leq M_t}. \tag{27}$$

In this case, $\sigma^2_{M_t,t} = 2S_2(x_t, x_t + M_t)$ and the condition in Theorem 5 can be seen to be equivalent to

$$\lim_{m \to +\infty} \frac{S_3(x_t, x_t + M_t)}{(S_2(x_t, x_t + M_t))^{3/2}} = 0. \tag{28}$$

That is, equation (28) is the only condition we have to check in order to apply the Invariance Principle to the terms in (27). Notice also (by a simple limiting argument) that if $S_2(1) < +\infty$, we can also take $M_t = +\infty$ in the above.

6 The scaling result for leadership

6.1 The general statement

Recalling Section 4.1, let f be a ELM function. There exists an x_0 such that $h(x) = xf'(x)/f(x) > 1/2$ for all $x \geq x_0$, which means that for such large x

$$q_0(x) \equiv \sqrt{\frac{x}{4h(x)-2}} \quad (x \geq x_0) \tag{29}$$

is a well-defined, positive function. Our generalization of Theorem 2 shows that the quantity $q_0(t)$ plays the same role that as the map $t \mapsto \sqrt{t/(4p-2)}$ in the specific case $f(x) = x^p$, $p > 1/2$. That is, in order to bias a balls in bins process started with t balls towards leadership by a given bin, the difference between the initial numbers of balls should be $\Theta\left(q_0(t)\right)$.

Theorem 6 *Let f be a ELM function and define q_0 as above. Let $\lambda \in \mathbf{R}$ be a constant, and assume that $q = q(n)$ ($n \in \mathbf{N}$) is such that*

- $t/2 \pm \lambda\, q(n) \in \mathbf{N}$ *for all* $n \in \mathbf{N}$*;*
- $q(n) \sim q_0(n)$ *for* $n \gg 1$*.*

Now consider the 2-bin balls-in-bins process started from initial state

$$(x(t), y(t)) = \left(\frac{t}{2} + \lambda\, q(t), \frac{t}{2} - \lambda\, q(t)\right) \in \mathbf{N}^2.$$

Then

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x(t),y(t))}\left(\mathsf{ELead}_1\right) = \Phi(\lambda), \tag{30}$$

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x(t),y(t))}\left(\mathsf{Lead}_1\right) = \Gamma(\lambda). \tag{31}$$

Table 1 presents estimates of $q_0(n)$ for n large, for several choices of feedback functions in the ELM regime. In particular, the case $f(x) = x^p$ ($p > 1/2$) of the Theorem implies Theorem 2. The remainder of this section contains the proof of the general result.

Proof: [of Theorem 6] We start by discussing how one can write the event ELead_1 and lead in terms of the exponential embedding. We only prove the result for $\lambda > 0$: the case $\lambda = 0$ is a simple extension, and the case $\lambda < 0$ reduces to the one we discuss below.

Let $(x, y) \in \mathbf{N}^B$ be the (for the time being arbitrary) initial conditions with $x > y$. The event ELead_1 holds whenever bin 1 reaches reach level M before bin 2 does for all large

$f(x) =$	$h(x) \sim$	$q_0(x) \sim$	conditions
$x^p \ln^q(x+e-1)$	p	$\sqrt{\frac{x}{4p-2}}$	$p > 1/2,\ q \in \mathbf{R}$
$x^{q \ln^\alpha(x)}$	$(\alpha+1) q \ln^\alpha(x)$	$\sqrt{\frac{x}{4(\alpha+1) q \ln^\alpha(x)}}$	$p > 1/2,\ q, \alpha > 0$
e^{x^p}	$p x^p$	$\frac{x^{\frac{1-p}{2}}}{2\sqrt{p}}$	$0 < p < 1/4$

Table 1: Some ELM feedback functions and their corresponding h and q_0. The last condition describes the conditions on the parameters p, q, α under which each f is indeed ELM.

enough M. This requires that the time it takes for bin 1 to reach level M in the continuous-time process is smaller than the corresponding time for bin 2. In the exponential embedding, this corresponds to

$$\exists M_0 \, \forall M \geq M_0 \ \sum_{j=x}^{M} X(1,j) < \sum_{j=y}^{M} X(2,j).$$

The above event can be rewritten as

$$\exists M_0 \, \forall M \geq M_0 \ \sum_{j=x}^{M} (X(1,j) - X(2,j)) < \sum_{j=y}^{x-1} X(1,j).$$

As noted in the proof sketch for Theorem 1 in Section 3.2, $\sum_{j=x}^{M} (X(1,j) - X(2,j))$ converges as $M \to +\infty$. If follows that, except for a null event, the above holds if and only if

$$\sum_{j=x}^{+\infty} (X(1,j) - X(2,j)) < \sum_{j=y}^{x-1} X(2,j).$$

Thus we deduce that

$$\mathbf{Pr}_{(x,y)}(\mathsf{ELead}_1) = \mathbf{Pr}\left(\sum_{j=x}^{+\infty} (X(1,j) - X(2,j)) < \sum_{j=y}^{x-1} X(2,j) \right). \tag{32}$$

What about the probability of Lead_1? Using the above notation, Lead_1 holds if for *all* $M \geq x$ bin 1 reaches level M before bin 2 does. That corresponds to

$$\forall M \geq x \ \sum_{j=x}^{M} X(1,j) < \sum_{j=y}^{M} X(2,j),$$

or

$$\forall M \geq x_{\ell_i} \ \sum_{j=x}^{M} (X(1,j) - X(2,j)) < \sum_{j=y}^{x-1} X(2,j).$$

It follows that

$$\mathbf{Pr}_{(x,y)}(\mathsf{Lead}_1) = \mathbf{Pr}\left(\sup_{M\geq x}\sum_{j=x}^{M}(X(1,j)-X(2,j)) < \sum_{j=y}^{x-1}X(2,j)\right). \tag{33}$$

Recall now the choice $x = x(t) = t/2 + \lambda q(t)$ and $y = y(t) = t/2 - \lambda q(t)$, where

$$q(n) \sim \sqrt{\frac{n}{4h(n)-2}} \text{ for } n \gg 1. \tag{34}$$

As discussed in Section 5.2, $\{X(1, x(t)+n-1) - X(2, x(t)+n-1)\}_{t\in\mathbf{N}, n\in\mathbf{N}}$ is a doubly-infinite array of centered, square-integrable random variables with

$$\sigma_{n,t}^2 \equiv \sum_{j=1}^{t}\mathbf{Var}\left(X(1,x(t)+n-1) - X(2,x(t)+n-1)\right) = 2S_2(x(t), x(t)+n) \ \ (1\leq n\leq +\infty).$$

One can construct a random continuous path $\Xi_t(\cdot)$ defined by setting

$$\Xi_t\left(\frac{S_2(x(t), x(t)+n)}{S_2(x(t))}\right) = \frac{\sum_{j=x}^{x(t)+n-1}X(1,j)-X(2,j)}{\sqrt{2S_2(x)}}, (1\leq n\leq +\infty) \tag{35}$$

and completing the remaining values by linear interpolation. As discussed in Section 5.2, the fact that we have infinite terms here poses no problems. Moreover, it is easy to check that

$$\sup_{0\leq s\leq 1}\Xi_t(s) = \frac{1}{\sqrt{2S_2(x)}}\sup_{M\geq x}\sum_{j=x}^{M}X(1,j)-X(2,j), \tag{36}$$

$$\Xi_t(1) = \frac{1}{\sqrt{2S_2(x)}}\sum_{j=x}^{+\infty}X(1,j)-X(2,j). \tag{37}$$

Thus one can rewrite

$$\mathbf{Pr}_{(x,y)}(\mathsf{ELead}_1) = \mathbf{Pr}\left(\Xi_t(1) < \frac{\sum_{j=y}^{x-1}X(2,j)}{\sqrt{2S_2(x)}}\right), \tag{38}$$

$$\mathbf{Pr}_{(x,y)}(\mathsf{Lead}_1) = \mathbf{Pr}\left(\sup_{0\leq s\leq 1}\Xi_t(s) < \frac{\sum_{j=y}^{x-1}X(2,j)}{\sqrt{2S_2(x)}}\right). \tag{39}$$

Suppose we show that as $t\to+\infty$

$$\Xi_t(\cdot) \to^w \text{ a standard Brownian Motion } B(\cdot), \tag{40}$$

$$\frac{\sum_{j=y}^{x-1}X(2,j)}{\sqrt{2S_2(x)}} \to^w \lambda. \tag{41}$$

Since $\sum_{j=y}^{x-1} X(2,j)$ is independent of $\Xi_t(\cdot)$, this means that

$$(\Xi_t(\cdot), \sum_{j=y}^{x-1} X(2,j)) \to^w (B(\cdot), \lambda),$$

which implies

$$\mathbf{Pr}_{(x,y)}(\mathsf{ELead}_1) \quad \to \quad \mathbf{Pr}(B(1) < \lambda), \tag{42}$$

$$\mathbf{Pr}_{(x,y)}(\mathsf{Lead}_1) \quad \to \quad \mathbf{Pr}\left(\sup_{0 \le s \le 1} B(s) < \lambda\right). \tag{43}$$

These probabilities can be evaluated via standard formulae for Brownian motion in Section 5.1, yielding the final result. We thus concentrate on proving equations (40) and (41).

Proof of (40). By Section 5.2, it suffices to show that $S_3(x(t)) \ll S_2(x(t))^{3/2}$. But this follows directly from the formulae in Lemma 1, the fact that $x = x(t) \to +\infty$, and the assumption that $h(x) \ll \sqrt{x}$:

$$S_3(x) \sim \frac{x}{(3h(x)-1)f(x)} \ll \left(\frac{x}{(2h(x)-1)f^2(x)}\right)^{3/2} = S_2(x)^{3/2}.$$

Proof of (41). Let us first establish a few facts about f, x, y and q.

1. *For $t \gg 1$, $q(t) \sim f(t/2)\sqrt{S_2(t/2)/2} = O(t)$.* Indeed, for t large, Lemma 1 implies that $S_2(t/2) \sim t/(4h(t/2)-2)f(t/2)^2$, and, because of (19) in Definition 1, $h(t/2) \sim h(t)$. Moreover, since $\liminf_{n\to+\infty} h(n) > 1/2$, $q(t) = \sqrt{t/\Omega(1)} = O(\sqrt{t})$.

2. *For $t \gg 1$, $f(x) = f(t/2 + \lambda q(t)) \sim f(t/2)$.* In this case

$$\begin{aligned} \left|\ln \frac{f(x)}{f\left(\frac{t}{2}\right)}\right| &= \left|\int_{\frac{t}{2}}^{\frac{t}{2}+\lambda q(t)} (\ln f(u))'\, du\right| \qquad (44) \\ &= \left|\int_{\frac{t}{2}}^{\frac{t}{2}+\lambda q(t)} \frac{h(u)}{u}\, du\right| \\ &\le \sup_{\frac{t}{2} \le u \le \frac{t}{2}+\lambda q(t)} |h(u)| \ln\left(1 + \frac{2q(t)}{t}\right). \end{aligned}$$

   Now notice that equation (19) implies that the sup of $h(u)$ above is $\sim h(t/2) \sim h(t)$. Moreover, $q(t)/t = \sqrt{1/t(2h(t)-1)} \ll 1$, since $\liminf_{n\to+\infty} h(n) > 1/2$. Therefore,

$$\left|\ln \frac{f(x)}{f\left(\frac{t}{2}\right)}\right| = O\left(\frac{h(t)q(t)}{t}\right) = O\left(\frac{1}{\sqrt{t}}\right) = o(1). \tag{45}$$

3. *For $t \gg 1$, $f(y) \sim f(t/2)$.* The proof is almost identical to the one above.

4. *For $t \gg 1$ and $r \geq 2$, $S_r(x), S_r(y) \sim S_r(t/2)$.* Indeed, because $q(t) = O\left(\sqrt{t}\right)$, $x, y \to +\infty$,, and the formulae in Lemma 1 apply. Thus $S_2(x) \sim x/[(rh(x)-1)f(x)^2]$, and by 1. and 3., $x \sim t/2$, $h(x) \sim h(t/2)$ and $f(x) \sim f(t/2)$, which implies $S_2(x) \sim (t/2)/[(rh(t/2)-1)f(t/2)^2] \sim S_r(t/2)$. The same argument proves the desired result for $S_r(y)$.

We now apply the estimates to the problem at hand. For t large enough (so that x and y are also large), we can ensure that f is increasing on $[y, x]$, so that

$$\frac{x-y}{f(x)} \leq \mathbf{Ex}\left[\sum_{\ell=y}^{x-1} X(2,\ell)\right] = \sum_{j=y}^{x-1} \frac{1}{f(j)} \leq \frac{x-y}{f(y)}.$$

By items 2. and 3. above and the definition of x, y, this implies that

$$\mathbf{Ex}\left[\sum_{\ell=y}^{x-1} X(2,\ell)\right] \sim \frac{2\lambda q(t)}{f(t/2)} \quad (t \gg 1). \tag{46}$$

Similarly, one can show that

$$\mathbf{Var}\left(\sum_{\ell=y}^{x-1} X(2,\ell)\right) \sim \frac{2\lambda q(t)}{f(t/2)^2} \quad (t \gg 1). \tag{47}$$

Therefore, using 4.,

$$\mathbf{Ex}\left[\sum_{\ell=y}^{x-1} X(2,\ell)\right]^2 \geq (1-o(1))2\lambda\, q(t)\, \mathbf{Var}\left(\sum_{\ell=y}^{x-1} X(2,\ell)\right) \quad (t \gg 1). \tag{48}$$

By Chebyshev's Inequality, it follows that

$$\frac{\sum_{\ell=y}^{x-1} X(2,\ell)}{\left(\frac{2\lambda q(t)}{f(t/2)}\right)} \to^w 1 \text{ as } t \to +\infty.$$

Finally, since

$$\frac{2q(t)}{f(t/2)} \sim \sqrt{\frac{2t}{(2h(t)-1)f(t/2)^2}} \sim \sqrt{2S_2(x)},$$

we have

$$\frac{\sum_{\ell=y}^{x-1} X(2,\ell)}{\lambda\sqrt{2S_2(x)}} \to^w 1 \text{ as } t \to +\infty,$$

which is the desired result. $\square$

7 The almost balanced regime

This section proves our result on the overtaking time, a generalization of Theorem 3. To recapitulate: Theorem 1 tells us that, when the feedback function f satisfies

$$\sum_{j=1}^{+\infty} \frac{1}{f(j)^2} = +\infty, \tag{49}$$

each of the two bins will be the one with more balls infinitely many times. Our main interest in this chapter will be in determining *how long* it takes for bin 1 to have more balls than bin 2, given that the latter bin has more balls at the start. More specifically, assume the process $(I_1(\cdot), I_2(\cdot))$ is started from state

$$(\lceil \alpha t \rceil, t - \lceil \alpha t \rceil), \;\; t \gg 1 \text{ and } \alpha \in (0, 1/2) \text{ fixed.} \tag{50}$$

As in the introduction, let V be the *overtaking time* of the process: that is the first time when bin 1 has more balls than bin 2.

$$V \equiv \min\{v \in \mathbf{N} : I_1(v) > I_2(v)\}. \tag{51}$$

Under condition (49), this min exists and is finite with probability 1. We will be interested in describing the asymptotic distribution of V.

To express our main result, let us introduce two mappings.

$$\begin{aligned} F_{t,\alpha} \;\; : \;\; ((1-\alpha)t, +\infty) \;\; &\to \;\; \mathbf{R}^+ \\ u \;\; &\mapsto \;\; \frac{\int_{\alpha t}^{(1-\alpha)t} \frac{dx}{f(x)}}{\sqrt{\int_{(1-\alpha)t}^{u} \frac{dx}{f(x)^2}}}, \end{aligned} \tag{52}$$

$$G_{t,\alpha} \equiv \text{the inverse of } F_{t,\alpha}. \tag{53}$$

Notice that $\lim_{u \searrow (1-\alpha)t} F_{t,\alpha}(u) = +\infty$ and

$$\lim_{u \to +\infty} F_{t,\alpha}(u) = \frac{\int_{\alpha t}^{(1-\alpha)t} \frac{dx}{f(x)}}{\sqrt{\int_{(1-\alpha)t}^{+\infty} \frac{dx}{f(x)^2}}} = 0$$

because (as a consequence of $\sum_j^{+\infty} f(j)^{-2} = +\infty$) the denominator in the RHS is infinite. Thus $F_{t,\alpha}$ is a monotone-decreasing function whose range is $\mathbf{R}^+$, and $G_{t,\alpha}$ is not only well-defined, but monotone-decreasing as well.

Theorem 7 *Assume that f is a AB function (cf. Definition 1), and define $F_{t,\alpha}$, $G_{t,\alpha}$ as above. Let $V_{t,\alpha}$ be the random variable V defined above, conditioned on the initial state $(\lceil \alpha t \rceil, t - \lceil \alpha t \rceil)$ of the balls-in-bins process. Then, as $t \to +\infty$,*

$$\forall \lambda \in \mathbf{R}^+, \lim_{t \to +\infty} \mathbf{Pr}\left(V_{t,\alpha} \geq 2G_{t,\alpha}(\lambda) - (t+1)\right) = \Gamma(\lambda), \tag{54}$$

(where $\Gamma(\cdot)$ is defined as in Theorem 2), or equivalently,

$$F_{t,\alpha}\left(\frac{V_{t,\alpha} + (t+1)}{2}\right) \to^w |N|, \tag{55}$$

for a standard Gaussian random variable N. The latter expression makes sense because $V_{t,\alpha} \geq (1-\alpha)t$, so

$$\frac{V_{t,\alpha} + (t+1)}{2} \geq \left(1 - \frac{\alpha}{2}\right) t > (1-\alpha)t.$$

This result is quite general, but applying it to a specific situation requires a calculation. We do this for the case $f(x) = x^p$ below, and then prove Theorem 7 below.

7.1 Proof of the special case

Claim 1 *Theorem 7 implies Theorem 3.*

Proof: One way of interpreting Theorem 7 in the $f(x) = x^p$ case is by saying that

$$V_{t,\alpha} \equiv 2G_{t,\alpha}(U_{t,\alpha,p}) - (t+1) \text{ with probability } \to 1,$$

where $U_{t,\alpha,p} \to^w |N|$ as $t \to +\infty$. Thus the Corollary follows from providing a formula for $G_{t,\alpha}$. We will first assume that $0 < p < 1/2$, in which case

$$\begin{aligned} e(t,\alpha) &\equiv \int_{\alpha t}^{(1-\alpha)t} \frac{dx}{f(x)} &(56)\\ &= [(1-\alpha)^{1-p} - \alpha^{1-p}]\frac{t^{1-p}}{1-p}, &(57)\end{aligned}$$

and for all $u > (1-\alpha)t$

$$\begin{aligned} f(u,t,\alpha) &\equiv \int_{(1-\alpha)t}^{u} \frac{dx}{f(x)^2} &(58)\\ &= \frac{u^{1-2p} - [(1-\alpha)t]^{1-2p}}{1-2p}. &(59)\end{aligned}$$

Then, for $u \geq (1-\alpha)t$,

$$F_{t,\alpha}(u) = \frac{e(t,\alpha)}{\sqrt{f(u,t,\alpha)}} \tag{60}$$

$$= \frac{[(1-\alpha)^{1-p} - \alpha^{1-p}]\frac{t^{1-p}}{1-p}}{\sqrt{\frac{u^{1-2p}-[(1-\alpha)t]^{1-2p}}{1-2p}}} \tag{61}$$

To compute $G_{t,\alpha}(\lambda)$ for some $\lambda \in \mathbf{R}^+$, we must solve the equation $F_{t,\alpha}(G_{t,\alpha}(\lambda)) = \lambda$, which corresponds to

$$G_{t,\alpha}(\lambda)^{1-2p} = [(1-\alpha)t]^{1-2p} + \frac{(1-2p)}{\lambda^2}\left\{[(1-\alpha)^{1-p} - \alpha^{1-p}]\frac{t^{1-p}}{1-p}\right\}^2. \tag{62}$$

Therefore,

$$G_{t,\alpha}(\lambda) = \left\{\frac{(1-2p)}{(1-p)^2}[(1-\alpha)^{1-p} - \alpha^{1-p}]^2 + \frac{(1-\alpha)^{1-2p}\lambda^2}{t}\right\}^{\frac{1}{1-2p}} \frac{t^{1+\frac{1}{1-2p}}}{\lambda^{1+\frac{2p}{1-2p}}}, \tag{63}$$

and the result for $0 < p < 1/2$ follows.

For $p = 1/2$, the above formula for $e(t,\alpha)$ still applies, but

$$f(u,t,\alpha) \equiv \int_{(1-\alpha)t}^{u} \frac{dx}{f(x)^2} \tag{64}$$

$$= \ln \frac{u}{(1-\alpha)t}, \tag{65}$$

and thus

$$G_{t,\alpha}(\lambda) = (1-\alpha)\, t \,\exp\left\{\frac{e(t,\alpha)^2}{\lambda^2}\right\} \tag{66}$$

$$= (1-\alpha)\, t \,\exp\left\{4[1-2\sqrt{\alpha(1-\alpha)}]\,\frac{t}{\lambda^2}\right\}. \tag{67}$$

This finishes the proof. $\square$

7.2 Proof of Theorem 7

Proof: Define the *overtaking number* $N_{t,\alpha}$ to be $I_1(V_{t,\alpha})$, i.e. the number $I_1(v)$ of balls in bin 1 at the first time v when $I_1(v) > I_2(v)$, under initial conditions $(\lceil \alpha t \rceil, t - \lceil \alpha t \rceil)$. It follows from this definition that at time $v' = V_{t,\alpha} - 1$, $I_1(v') = I_2(v') = N_{t,\alpha} - 1$. Since $I_1(0) + I_2(0) = t$, this means that

$$\begin{aligned} V_{t,\alpha} - 1 + t &= I_1(v') + I_2(v') \\ &= \text{total \# of balls at time } v' \\ &= 2N_{t,\alpha} - 2 \\ \Rightarrow V_{t,\alpha} &= 2N_{t,\alpha} - (t+1). \end{aligned} \tag{68}$$

Thus results about the distribution of $N_{t,\alpha}$ translate immediately into results about $V_{t,\alpha}$. Since $N_{t,\alpha}$ is easier to analyze via our techniques, we shall spend most of our time considering this quantity, returning to the more significant $V_{t,\alpha}$ at the end of the proof.

We begin by showing that, in terms of the exponential embedding random variables,

$$\forall M \in \mathbf{N} \;\; \{N_{t,\alpha} \geq M\} = \left\{ \sup_{y \leq m \leq M-1} \sum_{j=y}^{m-1} (X(2,j) - X(1,j)) \leq \sum_{\ell=x}^{y-1} X(1,\ell) \right\}, \tag{69}$$

where $x \equiv \lceil \alpha t \rceil$ (respectively $y \equiv t - \lceil \alpha t \rceil$) is the initial number of balls in bin 1 (resp. 2). Indeed, $N_{t,\alpha} \geq M$ occurs if and only if for all $y \leq m \leq M-1$, the time it takes for bin 2 to receive its mth ball (which is $\sum_{j=y}^{m-1} X(2,j)$) is smaller than or equal to the time it takes for bin 1 to receive its mth balls (which is $\sum_{\ell=x}^{m-1} X(1,\ell)$). Symbolically,

$$\forall M \in \mathbf{N} \;\; \{N_{t,\alpha} \geq M\} = \left\{ \forall y \leq m \leq M-1, \; \sum_{j=y}^{m-1} X(2,j) \leq \sum_{\ell=x}^{m-1} X(1,\ell) \right\},$$

from which (69) follows.

We now wish to estimate the probability of the event at the RHS of (69). We begin by looking at

$$\sup_{y \leq m \leq M-1} \sum_{j=y}^{m-1} (X(2,j) - X(1,j)). \tag{70}$$

In particular, let us assume that

$$M \equiv \lfloor G_{t,\alpha}(\lambda) \rfloor \tag{71}$$

for some $\lambda \in \mathbf{R}^+$, so that

$$\{N_{t,\alpha} \geq M\} = \{N_{t,\alpha} \geq G_{t,\alpha}(\lambda)\}. \tag{72}$$

We wish to apply the Invariance Principle to the random variables

$$\{\xi_{n,t} \equiv X(2, y(t) + n - 1) - X(1, y(t) + n - 1)\}_{t \in \mathbf{N}, 1 \leq n \leq M - y(t) - 1}.$$

Note that

$$\sigma_{n,t}^2 \equiv \sum_{j=1}^{n} \mathbf{Var}\left(X(2, y(t)+n-1) - X(1, y(t)+n-1)\right) = 2S_2(y(t), y(t)+n).$$

Following Section 5.2, we build the random path $\Xi_t(\cdot)$ that linearly interpolates the values

$$\Xi_t\left(\frac{S_2(y(t), y(t)+n)}{S_2(y(t), M-1)}\right) \equiv \frac{\sum_{j=1}^{n} X(2, y(t)+j-1) - X(1, y(t)+j-1)}{\sqrt{2S_2(y, M-1)}},\ 1 \le n \le M - y(t).$$

As in the previous proof, we have

$$\sup_{0\le s\le 1} \Xi_t(s) = \frac{\sup_{y\le m\le M-1} \sum_{j=y}^{m-1} X(2, y(t)+j-1) - X(1, y(t)+j-1)}{\sqrt{2S_2(y, M-1)}}.$$

Hence

$$\mathbf{Pr}\left(N_{t,\alpha} \ge M\right) = \mathbf{Pr}\left(\sup_{0\le s\le 1} \Xi_t(s) \le \frac{\sum_{j=x}^{y-1} X(1,j)}{\sqrt{2S_2(y, M-1)}}\right).$$

We will eventually prove that as $t \to +\infty$,

$$\sum_{j=x}^{y-1} X(1,j) \quad \to^{w} \quad \lambda, \tag{73}$$

$$\Xi_t(\cdot) \quad \to^{w} \quad \text{a standard Brownian Motion } B(\cdot). \tag{74}$$

It follows from this and the independence of $\Xi_t(\cdot)$ $\sum_{j=x}^{y-1} X(1,j)$ that

$$\lim_{t\to+\infty} \mathbf{Pr}\left(N_{t,\alpha} \ge M\right) = \mathbf{Pr}\left(\sup_{0\le s\le 1} B(s) \le \lambda\right) = \Gamma(\lambda), \tag{75}$$

which is the desired result. Thus we concentrate on proving (74) and (73).

Proof of (73). The expectation of $\sum_{\ell=x}^{y-1} X(1,\ell)$ can be estimated as follows.

$$\mathbf{Ex}\left[\sum_{\ell=x}^{y-1} X(1,\ell)\right] = S_1(\lceil \alpha t\rceil, t - \lceil \alpha t\rceil) \tag{76}$$

$$\sim \int_{\alpha t}^{(1-\alpha)t} \frac{ds}{f(s)} \tag{77}$$

$$= \Omega\left(\frac{t}{f(\alpha t)}\right) \gg 1. \tag{78}$$

Indeed, (77) follows from the fact that $\sum_{j\geq 1} f(j)^{-1} = +\infty$, and (78) follows from the assumption $h(s) \leq 1/2$ for s large, which means that $f((1-\alpha)t) = O\left(f(\alpha t)\right) = O\left(\sqrt{t}\right)$. The variance of the is

$$\mathbf{Var}\left(\sum_{\ell=x}^{y-1} X(1,\ell)\right) = S_2(\lceil \alpha t\rceil, t - \lceil \alpha t\rceil) = O\left(\frac{t}{f(\alpha t)^2}\right) \ll \mathbf{Ex}\left[\sum_{\ell=x}^{y-1} X(1,\ell)\right]^2 \quad (t \gg 1)$$

because for s large, $h(s) \geq 0$ and thus f is decreasing. By Chebyshev's Inequality, it follows that

$$\frac{\sum_{\ell=x}^{y-1} X(1,\ell)}{\int_{\alpha t}^{(1-\alpha)t} \frac{ds}{f(s)}} \to^{w} 1. \tag{79}$$

On the other hand, notice that if $M' \equiv G_{t,\alpha}(\lambda)$ satisfies

$$F_{t,\alpha}(M') = \frac{\int_{\alpha t}^{(1-\alpha)t} \frac{ds}{f(s)}}{\sqrt{2\int_{(1-\alpha)t}^{M'} \frac{ds}{f(s)^2}}} = \lambda. \tag{80}$$

We wish to show that

$$F_{t,\alpha}(M) \sim \lambda; \tag{81}$$

this will follow from

$$\int_{(1-\alpha)t}^{M'} \frac{ds}{f(s)^2} \sim \int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2}.$$

Since $M = \lfloor M' \rfloor$, $M \leq M' \leq M+1$ and

$$\left|\int_{(1-\alpha)t}^{M'} \frac{ds}{f(s)^2} - \int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2}\right| \leq \frac{1}{f(M')^2}.$$

This follows from the fact that $h(s) > 0$ (and thus f is increasing) for s large enough. On the other hand, we must have $M' - (1-\alpha)t \gg 1$, as for any constant C

$$\int_{(1-\alpha)t}^{(1-\alpha)t+C} \frac{ds}{f(s)^2} \leq \frac{C}{f((1-\alpha)t)} \ll 1 \ll \left(\int_{\alpha t}^{(1-\alpha)t} \frac{ds}{f(s)}\right)^2,$$

contradicting (80). Thus

$$\int_{(1-\alpha)t}^{M'} \frac{ds}{f(s)^2} \geq \frac{M' - (1-\alpha)t}{f(M')^2} \gg \left|\int_{(1-\alpha)t}^{M'} \frac{ds}{f(s)^2} - \int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2}\right|,$$

which proves (81). In particular, it follows from (79) that

$$\frac{\sum_{j=x}^{y-1} X(1,j)}{\int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2}} \to^{w} \lambda. \tag{82}$$

Finally, since by (78) the numerator of

$$F_{t,\alpha}(M) = \frac{\int_{\alpha t}^{(1\alpha)t} \frac{ds}{f(s)^2}}{\sqrt{2\int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2}}}$$

diverges, we deduce that $\int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2} \gg 1$, which implies that

$$\int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2} \sim S_2(y, M-1).$$

Plugging this into (82) yields (73).

Proof of (74). We have already shown above that $S_2(y, M-1) \sim \int_{(1-\alpha)t}^{M} \frac{ds}{f(s)^2} \gg 1$. Since f is bounded below,

$$S_3(y, M-1) = O\left(S_2(y, M-1)\right) \ll S_2(y, M-1)^{3/2}.$$

Thus the condition for the Invariance Principlie in Section 5.2 is satisfied, and this finishes the proof. □

8 Extensions and open problems

Our Theorem 6 admits an extension to the case of a general number $B \geq 2$ of bins.

Theorem 8 *Let f be a ELM function, $B \geq 2$ and $\lambda_i \in \mathbf{R}$ ($i \in B$) be constants, and assume that $q = q(n)$ ($n \in \mathbf{N}$) is such that*

- *$t/B + \lambda_i\, q(n) \in \mathbf{N}$ for all $n \in \mathbf{N}$ and $i \in [B]$;*
- *$q(n) \sim q_0(n)$ for $n \gg 1$ (q_0 is defined in (29)).*

Now consider the B-bin balls-in-bins process started from initial state

$$(x_i(t))_{i=1}^{B} = \left(\frac{t}{B} + \lambda_i\, q(t)\right)_{i=1}^{B} \in \mathbf{N}^B.$$

Then

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x_i(t))_{i=1}^B}(\mathsf{ELead}_1) = \mathbf{Pr}\left(\forall 2\leq i\leq B\,,\, \frac{N_1-N_i}{\sqrt{2}} < (\lambda_1-\lambda_i)\right)$$

$$\lim_{t\to+\infty} \mathbf{Pr}_{(x_i(t))_{i=1}^B}(\mathsf{Lead}_1) = \mathbf{Pr}\left(\forall 2\leq i\leq B\,,\, \sup_{0\leq t\leq 1}\frac{(B_1(t)-B_i(t))}{\sqrt{2}} < (\lambda_1-\lambda_i)\right),$$

where $\{B_i(\cdot)\}_{i=1}^B$ *(*$\{N_i\}_{i=1}^B$*) are i.i.d. standard Brownian Motions (resp. Gaussians).*

The proof of this result, which we omit, follows essentially the same lines as that of Theorem 6. The only major differences is that the Invariance Principle is applied to the sequences

$$\{\xi_{n,t}^{(i)} = X(i, x_i(t)+n-1) - f(x_i(t)+n-1)^{-1}\}_{t\in\mathbf{N},n\in\mathbf{N}}\ (i\in[B])$$

and some extra care must be taken in looking at the maxima of the corresponding random continuous functions.

It is not entirely obvious how Theorem 7 on the almost balanced regime should or could be generalized to more than 2 bins. All we can say in general is that the addition of bins to the process will always make any overtaking take longer (to see this, notice that more bins just add more arrivals in the continuous-time exponential embedding between the start and overtaking times). In fact, there is a more basic question about such processes that we cannot answer; it was posed as a conjecture by Joel Spencer.

Conjecture 1 (Joel Spencer) *Consider a balls-in-bins process* $(I_m(i))_{i\in[B],m\geq 0}$ *with feedback function* f *that is in the almost-balanced regime (cf. Theorem 1). Then for all permutations* Π *of* $[B]$ *and all initial conditions, with probability* 1 *there are infinitely many* $m\geq 0$ *with* $I_m(\Pi(1)) < I_m(\Pi(2)) < \cdots < I_m(\Pi(B))$*. That is, all permutations possible of the bins occur infinitely often almost surely.*

Perhaps the techniques presented here could be used to settle this problem.

References

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002. Arxiv: `cond-mat/0106096`.

[2] Noga Alon and Joel Spencer. *The Probabilistic Method.* Wiley-Interscience Series in Discrete Mathematics. John Wiley and Sons, New York, second edition, 2000.

[3] Patrick Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Statistics. John Wiley and Sons, New York, second edition, 1999.

[4] Burgess Davis. Reinforced random walk. *Probability Theory and Related Fields*, 84(2):203–229, 1990.

[5] Eleni Drinea, Mihaela Enachescu, and Michael Mitzenmacher. Variations on Random Graph models of the Web. Harvard Technical Report TR-06-01, 2001.

[6] Eleni Drinea, Alan Frieze, and Michael Mitzenmacher. Balls in bins processes with feedback. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 308–315. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.

[7] Kostya Khanin and Raya Khanin. A probabilistic model for the establishment of neuron polarity. *Journal of Mathematical Biology*, 42(1):26–40, 2001.

[8] P.L. Krapivsky and Sidney L. Redner. Organization of growing random networks. *Physics Reviews E*, 63:066123, 2001. Available at Arxiv: `cond-mat/0011094`.

[9] Michael Mitzenmacher, Roberto Oliveira, and Joel Spencer. A scaling result for explosive processes. *Electronic Journal of Combinatorics*, 11(1):R31, 2004.

[10] Roberto Oliveira. *Preferential attachment*. PhD thesis, Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, 2004.

[11] Roberto Oliveira. The Onset of Dominance in Balls-in-Bins Processes with Feedback. Preprint. Arxiv: `math.PR/0510415`, 2005.

[12] Roberto Oliveira and Joel Spencer. Avoiding an iminent defeat in a balls-in-bins process with feedback. Manuscript, 2005.

[13] Joel Spencer and Nicholas Wormald. Explosive processes. Manuscript.