Discrete Mathematics and Theoretical Computer Science DMTCS vol. **10:3**, 2008, 171–196

Convergence of some leader election algorithms

Svante Janson[1] and Christian Lavault[2] and Guy Louchard[3]

[1] *Uppsala University, Department of Mathematics, PO Box 480, SE-751 06 Uppsala, Sweden.*
svante.janson@math.uu.se
[2] *LIPN (UMR CNRS 7030), Université Paris 13, 99, av. J.-B. Clément 93430 Villetaneuse, France.*
lavault@lipn.univ-paris13.fr
[3] *Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium.* louchard@ulb.ac.be

received February 8, 2008, revised August 19, 2008, December 11, 2008, accepted November 27, 2008.

We start with a set of n players. With some probability $P(n,k)$, we kill $n-k$ players; the other ones stay alive, and we repeat with them. What is the distribution of the number X_n of *phases* (or rounds) before getting only one player? We present a probabilistic analysis of this algorithm under some conditions on the probability distributions $P(n,k)$, including stochastic monotonicity and the assumption that roughly a fixed proportion α of the players survive in each round.

We prove a kind of convergence in distribution for $X_n - \log_{1/\alpha} n$; as in many other similar problems there are oscillations and no true limit distribution, but suitable subsequences converge, and there is an absolutely continuous random variable Z such that $d(X_n, \lceil Z + \log_{1/\alpha} n \rceil) \to 0$, where d is either the total variation distance or the Wasserstein distance.

Applications of the general result include the leader election algorithm where players are eliminated by independent coin tosses and a variation of the leader election algorithm proposed by W.R. Franklin 1982. We study the latter algorithm further, including numerical results.

Keywords: Leader election algorithms, convergence in distribution, oscillations

1 A general convergence theorem

We consider a general leader election algorithm of the following type: We are given some random procedure that, given any set of $n \geq 2$ individuals, eliminates some (but not all) individuals. If there is more that one survivor, we repeat the procedure with the set of survivors until only one (the winner) remains. We are interested in the (random) number X_n of rounds required if we start with n individuals. (We set $X_1 = 0$, and have $X_n \geq 1$ for $n \geq 2$.) We let N_k be the number of individuals remaining after round k; thus $X_n := \min\{k : N_k = 1\}$, where we start with $N_0 = n$. For convenience we may suppose that we continue with infinitely many rounds where nothing happens; thus N_k is defined for all $k \geq 0$ and $N_k = 1$ for all $k \geq X_n$.

We assume that the number Y_n of survivors of a set of n individuals has a distribution depending only on n. We have $1 \le Y_n \le n$; we allow the possibility that $Y_n = n$, but we assume $\mathbb{P}(Y_n = n) < 1$ for every $n \ge 2$, so that we will not get stuck before selecting a winner. We further assume that, given the number of remaining individuals at the start of a new round, the number of survivors is independent of the previous history. In other words, the sequence $(N_k)_0^\infty$ is a Markov chain on $\{1, 2, \ldots\}$, and X_n is the number of steps to absorption in 1. The transition probabilities of this Markov chain are, with $Y_1 = 1$,

$$P(i, j) := \mathbb{P}(Y_i = j) = \mathbb{P}(j \text{ survives of a set of } i). \tag{1.1}$$

Note that $P(i, j) = 0$ if $j > i$ and $P(i, i) < 1$, for $i > 1$. Conversely, any Markov chain on $\{1, 2, \ldots\}$ with such $P(i, j)$ can be regarded as a leader election algorithm in the generality just described.

We will in this paper treat leader election algorithms where, asymptotically, a fixed proportion is eliminated in each round. (Thus, we expect X_n to be of the order $\log n$.) More precisely, we assume the following for Y_n, where we also repeat the key assumptions above. (Here and below, $\log n$ should be interpreted as some fixed positive number when $n = 1$.)

Condition 1.1. *For every $n \ge 1$, Y_n is a random variable such that $1 \le Y_n \le n$, and $\mathbb{P}(Y_n = n) < 1$ for $n \ge 2$. Further:*

(i) *Y_n is stochastically increasing in n, i.e., $\mathbb{P}(Y_n \le k) \ge \mathbb{P}(Y_{n+1} \le k)$ for all $n \ge 1$ and $k \ge 1$. Equivalently, we may couple Y_n and Y_{n+1} such that $Y_n \le Y_{n+1}$.*

(ii) *For some constants $\alpha \in (0, 1)$ and $\varepsilon > 0$ and a sequence $\delta_n = O\big((\log n)^{-1-\varepsilon}\big)$,*

$$\mathbb{E}Y_{n+1} - \mathbb{E}Y_n = \alpha + \mathcal{O}(\delta_n). \tag{1.2}$$

(iii) *For some ε and δ_n as in* (ii)*,*

$$\mathbb{P}(|Y_n - \alpha n| > \delta_n n) = \mathcal{O}(n^{-2-\varepsilon}). \tag{1.3}$$

Note that

$$\mathbb{E}|Y_n - \alpha n|^p = \mathcal{O}(n^{p/2}) \tag{1.4}$$

for some $p > 4$ suffices for (iii), for a suitable choice of $\varepsilon > 0$ and δ_n (e.g., $\delta_n = n^{-\eta}$, $\eta > 0$ and ε small).

Remark 1.2. If (1.2) or (1.3) holds for some sequence (δ_n), it holds for every larger sequence (δ_n) too; similarly, if $\delta_n = O\big((\log n)^{-1-\varepsilon}\big)$ or (1.3) holds for some ε, it holds for every smaller ε too. Hence we may assume that (ii) and (iii) hold with the same $\varepsilon > 0$ and the same δ_n, and we may assume $\delta_n \ge (\log n)^{-1-\varepsilon}$. In particular, this implies that $\delta_k = \mathcal{O}(\delta_n)$ when $C^{-1}n \le k \le Cn$, for each constant C.

The behaviour of the election algorithm is given by the recursion $X_1 = 0$ and

$$X_n \overset{\mathrm{d}}{=} X_{Y_n} + 1, \qquad n \ge 2, \tag{1.5}$$

where we assume that $(X_i)_{i=1}^n$ and Y_n are independent. We state a general convergence theorem for leader election algorithms of this type.

We recall the definitions of the total variation distance d_{TV} and the Wasserstein distance d_{W} (also known as the Dudley, Fortet-Mourier or Kantorovich distance, or minimal L_1 distance); these are both metrics on spaces of probability distributions, but it is convenient to write also $d_{\mathrm{TV}}(X,Y) := d_{\mathrm{TV}}(\mu,\nu)$ and $d_{\mathrm{W}}(X,Y) := d_{\mathrm{W}}(\mu,\nu)$ for random variables X,Y with $X \sim \mu$ and $Y \sim \nu$.

The total variation distance d_{TV} between (the distributions of) two arbitrary random variables X and Y is defined by

$$d_{\mathrm{TV}}(X,Y) := \sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|. \tag{1.6}$$

For integer-valued random variables, as is the case in our theorem, this is easily seen to be equivalent to

$$d_{\mathrm{TV}}(X,Y) = \tfrac{1}{2} \sum_k |\mathbb{P}(X = k) - \mathbb{P}(Y = k)|. \tag{1.7}$$

Further, for any distributions μ and ν,

$$d_{\mathrm{TV}}(\mu,\nu) := \inf \{\mathbb{P}(X \neq Y) : X \sim \mu,\, Y \sim \nu\}\,; \tag{1.8}$$

the infimum is taken over all random vectors (X,Y) on a joint probability space with the given marginal distributions μ and ν. (In other words, over all couplings (X,Y) of μ and ν.) For integer-valued random variables, convergence in d_{TV} is equivalent to convergence in distribution, or equivalently, weak convergence of the corresponding distributions.

The Wasserstein distance d_{W} is defined only for probability distributions with finite expectation, and can be defined by, in analogy with (1.8),

$$d_{\mathrm{W}}(\mu,\nu) := \inf \{\mathbb{E}|X - Y| : X \sim \mu,\, Y \sim \nu\}\,. \tag{1.9}$$

There are several equivalent formulas. For example, for integer-valued random variables,

$$d_{\mathrm{TV}}(X,Y) = \sum_k |\mathbb{P}(X \le k) - \mathbb{P}(Y \le k)|. \tag{1.10}$$

It is immediate from (1.8) and (1.9) that for integer-valued random variables X and Y (but not in general),

$$d_{\mathrm{TV}}(X,Y) \le d_{\mathrm{W}}(X,Y). \tag{1.11}$$

It is well-known that d_{W} is a complete metric on the space of probability measures on $\mathbb{R}$ with finite expectation, and that convergence in d_{W} is equivalent to weak convergence plus convergence of the first absolute moment.

All unspecified limits in this paper are as $n \to \infty$.

Theorem 1.3. *Consider the leader election algorithm described above, with Y_n satisfying Condition 1.1. Then, there exists a distribution function F with bounded density function $f = F'$ such that*

$$\sup_{k \in \mathbb{Z}} |\mathbb{P}(X_n \le k) - F(k - \log_{1/\alpha} n)| \to 0 \tag{1.12}$$

or, equivalently, if $Z \sim F$,

$$d_{\mathrm{TV}}(X_n, \lceil Z + \log_{1/\alpha} n \rceil) \to 0. \tag{1.13}$$

More precisely, $d_W(X_n, \lceil Z + \log_{1/\alpha} n \rceil) \to 0$, *which is equivalent to*

$$\sum_{k\in\mathbb{Z}} |\mathbb{P}(X_n \le k) - F(k - \log_{1/\alpha} n)| \to 0. \tag{1.14}$$

As a consequence, defining $\Delta F(x) := F(x) - F(x-1)$,

$$\sup_{k\in\mathbb{Z}} |\mathbb{P}(X_n = k) - \Delta F(k - \log_{1/\alpha} n)| \to 0. \tag{1.15}$$

Furthermore,

$$\mathbb{E}X_n = \log_{1/\alpha} n + \phi(n) + o(1), \tag{1.16}$$

for a continuous function $\phi(t)$ *on* $(0,\infty)$ *which is periodic in* $\log_{1/\alpha} t$, *i.e.* $\phi(t) = \phi(\alpha t)$, *and locally Lipschitz.*

We thus do not have convergence in distribution as $n \to \infty$, but the usual type of oscillations with an asymptotic periodicity in $\log_{1/\alpha} n$ and convergence in distribution along subsequences such that the fractional part $\{\log_{1/\alpha} n\}$ converges. (This phenomenon is well-known for many other problems with integer-valued random variables, see for example [16, 12]; it happens frequently when the variance stays bounded.) This is illustrated in Figure 1.

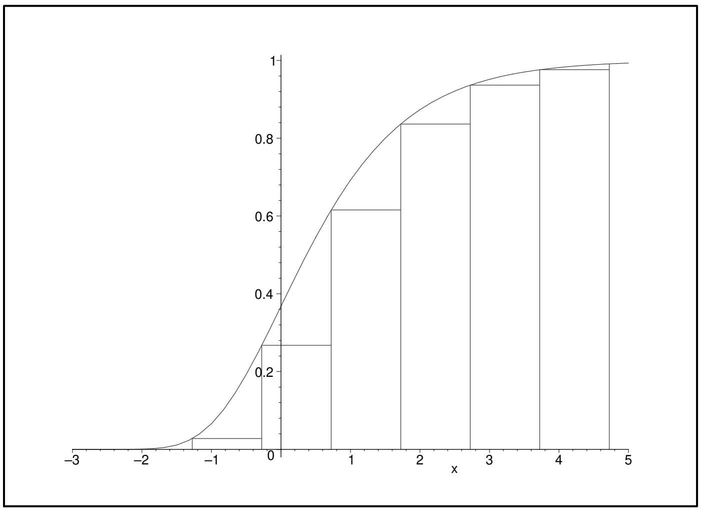


Fig. 1: Illustration of Theorem 1.3

Proof: We assume that δ_n are as in Remark 1.2.

Let $q := \sup_{n\ge 2} \mathbb{P}(Y_n = n)$. Since each $\mathbb{P}(Y_n = n) < 1$, and $\mathbb{P}(Y_n = n) \to 0$ by (iii), $q < 1$. Hence X_n is stochastically dominated by a sum of $n-1$ geometric $\mathrm{Ge}(1-q)$ random variables, and thus $\mathbb{E}X_n = \mathcal{O}(n)$. In particular, $\mathbb{E}X_n < \infty$ for every n.

Since the sequence (Y_n) is stochastically increasing, we may couple all Y_n such that $Y_1 \le Y_2 \le \dots$. If we consider starting our algorithm with different initial values, and use this coupling of (Y_n) in each round, we obtain a coupling of all X_n, $n \geq 1$, such that $X_{n+1} \geq X_n$ a.s. for every $n \geq 1$. We use these couplings of (Y_n) and (X_n) throughout the proof.

Let

$$\begin{aligned} x_n &:= \mathbb{E}X_n, \\ d_n &:= \mathbb{E}X_{n+1} - \mathbb{E}X_n = x_{n+1} - x_n, \\ b_n &:= \max_{1 \leq k \leq n} kd_k. \end{aligned}$$

We extend b_n to real arguments by the same formula; thus, $b_t = b_{\lfloor t \rfloor}$ for real $t \geq 1$.

By (1.5),

$$x_n = \mathbb{E}X_n = 1 + \mathbb{E}x_{Y_n}, \qquad n \geq 2.$$

Thus, for $n \geq 2$,

$$\begin{aligned} d_n &= \mathbb{E}(x_{Y_{n+1}} - x_{Y_n}) = \mathbb{E} \sum_{Y_n}^{Y_{n+1}-1} d_j = \mathbb{E} \sum_{j=1}^{n} d_j [\![Y_n \leq j < Y_{n+1}]\!] \\ &= \sum_{j=1}^{n} d_j \mathbb{P}(Y_n \leq j < Y_{n+1}). \end{aligned} \tag{1.17}$$

By (ii), $\mathbb{E}Y_{n+1} - \mathbb{E}Y_n \to \alpha$, and thus there exists n_0 such that if $n \geq n_0$ then

$$\sum_{j=1}^{n} \mathbb{P}(Y_n \leq j < Y_{n+1}) = \mathbb{E}Y_{n+1} - \mathbb{E}Y_n < 1.$$

Hence (1.17) implies, with $d_n^* = \max_{k \leq n} d_k$, for $n \geq n_0$,

$$d_n(1 - \mathbb{P}(Y_n \leq n < Y_{n+1})) \leq \sum_{j=1}^{n-1} \mathbb{P}(Y_n \leq j < Y_{n+1}) d_{n-1}^* \leq d_{n-1}^*(1 - \mathbb{P}(Y_n \leq n < Y_{n+1})),$$

and thus $d_n \leq d_{n-1}^*$ so $d_n^* = d_n \vee d_{n-1}^* = d_{n-1}^*$. Consequently, $d_n^* = d_{n_0}^* < \infty$, for all $n \geq n_0$. In other words, $d^* := \sup_n d_n < \infty$.

Let $\beta := (1 + \alpha)/2$; thus $\alpha < \beta < 1$. If n is large enough, so that $(\alpha + \delta_{n+1})(n+1) < \beta n$, then (1.17) yields, using (1.3) and (1.2),

$$\begin{aligned} d_n &\leq d^* \sum_{j < (\alpha - \delta_n)n} \mathbb{P}(Y_n \leq j) + \frac{1}{n(\alpha - \delta_n)} b_{\beta n} \sum_j \mathbb{P}(Y_n \leq j < Y_{n+1}) + d^* \sum_{j > \beta n} \mathbb{P}(Y_{n+1} > j) \\ &\leq d^* \mathcal{O}(nn^{-2-\varepsilon}) + \frac{1}{n(\alpha - \delta_n)} b_{\beta n} (\mathbb{E}Y_{n+1} - \mathbb{E}Y_n) \\ &= \mathcal{O}(n^{-1-\varepsilon}) + \frac{1}{n} \frac{\alpha + \mathcal{O}(\delta_n)}{\alpha - \delta_n} b_{\beta n}. \end{aligned}$$

Thus

$$nd_n \leq (1 + \mathcal{O}(\delta_n)) b_{\beta n} + \mathcal{O}(n^{-\varepsilon}). \tag{1.18}$$

Replace n by k and take the supremum over all k such that $\beta n < k \leq n$. Since b_k is increasing, and by our simplifying assumptions in Remark 1.2, this yields

$$b_n \leq (1 + \mathcal{O}(\delta_n))b_{\beta n} + \mathcal{O}(n^{-\varepsilon}) = \bigl(1 + \mathcal{O}(\delta_n)\bigr)b_{\beta n}.$$

It follows by induction over m that if $(1/\beta)^m \leq n < (1/\beta)^{m+1}$, then

$$b_n \leq C_1 \prod_{j=1}^{m} \Bigl(1 + \frac{C_2}{j^{1+\varepsilon}}\Bigr)$$

and thus $b_n = \mathcal{O}(1)$. In other words, we have shown

$$d_n = \mathcal{O}(1/n). \tag{1.19}$$

We now use the Wasserstein distance d_{W}. Since $X_{n+1} \geq X_n$ a.s., it is easily seen by (1.9) that $d_{\mathrm{W}}(X_n, X_{n+1}) = \mathbb{E}(X_{n+1} - X_n) = d_n$. Thus, if $m \leq n$, by (1.19),

$$d_{\mathrm{W}}(X_n, X_m) \leq \sum_{k=m}^{n-1} d_{\mathrm{W}}(X_k, X_{k+1}) = \sum_{k=m}^{n-1} d_k = \mathcal{O}\left(\frac{n-m}{m}\right),$$

and thus, for all n and m,

$$d_{\mathrm{W}}(X_n, X_m) = \mathcal{O}\left(\frac{|n-m|}{n \wedge m}\right). \tag{1.20}$$

Note also that (iii) implies

$$\mathbb{E}|Y_n - \alpha n| \leq \delta_n n + \mathcal{O}(n^{-1-\varepsilon}) = \mathcal{O}(n\delta_n). \tag{1.21}$$

Define

$$\widetilde{X}_t := X_{\lfloor t \rfloor} - \log_{1/\alpha} t, \qquad t \geq 1. \tag{1.22}$$

Then, for $t \geq 2/\alpha$, using (1.5), (1.20), (1.21), (1.3), and $1 \leq Y_{\lfloor t \rfloor} \leq t$,

$$\begin{aligned}
d_{\mathrm{W}}(\widetilde{X}_t, \widetilde{X}_{\alpha t}) &= d_{\mathrm{W}}\bigl(X_{\lfloor t \rfloor} - \log_{1/\alpha}(t), X_{\lfloor \alpha t \rfloor} - \log_{1/\alpha}(\alpha t)\bigr) \\
&= d_{\mathrm{W}}\bigl(X_{\lfloor t \rfloor} - 1, X_{\lfloor \alpha t \rfloor}\bigr) \leq \mathbb{E} d_{\mathrm{W}}\bigl(X_{Y_{\lfloor t \rfloor}}, X_{\lfloor \alpha t \rfloor}\bigr) \\
&\leq C_3 \mathbb{E}\left(\frac{|Y_{\lfloor t \rfloor} - \lfloor \alpha t \rfloor|}{Y_{\lfloor t \rfloor} \wedge \lfloor \alpha t \rfloor}\right) \\
&\leq C_3 \mathbb{E}\left(\frac{|Y_{\lfloor t \rfloor} - \lfloor \alpha t \rfloor|}{\alpha t/2} + t[\![Y_{\lfloor t \rfloor} < \alpha t/2]\!]\right) \\
&= \mathcal{O}\left(t^{-1}\mathbb{E}|Y_{\lfloor t \rfloor} - \lfloor \alpha t \rfloor|\right) + \mathcal{O}\left(t\mathbb{P}\left(Y_{\lfloor t \rfloor} < \alpha t/2\right)\right) \\
&= \mathcal{O}(\delta_{\lfloor t \rfloor}) + \mathcal{O}(t \cdot t^{-2-\varepsilon}) = \mathcal{O}\left(\log^{-1-\varepsilon} t\right).
\end{aligned}$$

Hence, for any $t \geq 2/\alpha$,

$$\sum_{j=0}^{\infty} d_{\mathrm{W}}(\widetilde{X}_{\alpha^{-j}t}, \widetilde{X}_{\alpha^{-j-1}t}) = \mathcal{O}\left(\log^{-\varepsilon} t\right) < \infty. \tag{1.23}$$

Since d_{W} is a complete metric, thus there exists for every $t > 0$ a limiting distribution $\mu(t)$, such that if $Z(t) \sim \mu(t)$, then

$$d_{\mathrm{W}}\big(\widetilde{X}_{\alpha^{-j}t}, Z(t)\big) \to 0 \quad \text{as} \quad j \to \infty. \tag{1.24}$$

In particular,

$$\widetilde{X}_{\alpha^{-j}t} \xrightarrow{\mathrm{d}} Z(t) \quad \text{as} \quad j \to \infty. \tag{1.25}$$

(We find it more convenient to use the random variable $Z(t)$ than its distribution $\mu(t)$.) Clearly, $Z(\alpha t) \overset{d}{=} Z(t)$, so the distribution $\mu(t)$ is a periodic function of $\log_{1/\alpha} t$. Hence, (1.24) can also be written, adding the explicit estimate obtained from (1.23),

$$d_{\mathrm{W}}\big(\widetilde{X}_t, Z(t)\big) = \mathcal{O}\left(\log^{-\varepsilon} t\right) \to 0 \quad \text{as} \quad t \to \infty. \tag{1.26}$$

Note further that, for $\gamma \geq 1$, by (1.22) and (1.20),

$$\begin{aligned} d_{\mathrm{W}}(\widetilde{X}_t, \widetilde{X}_{\gamma t}) &\leq d_{\mathrm{W}}(X_{\lfloor t \rfloor}, X_{\lfloor \gamma t \rfloor}) + |\log_{1/\alpha} t - \log_{1/\alpha}(\gamma t)| = \mathcal{O}\left(\frac{\lfloor \gamma t \rfloor - \lfloor t \rfloor}{t}\right) + \log_{1/\alpha} \gamma \\ &= \mathcal{O}(\gamma - 1 + 1/t). \end{aligned}$$

Replacing t by $\alpha^{-j}t$ and letting $j \to \infty$, it follows from (1.24) that, for all $t > 0$ and $\gamma \geq 1$,

$$d_{\mathrm{W}}(Z(t), Z(\gamma t)) = \mathcal{O}(\gamma - 1). \tag{1.27}$$

Consequently, $t \to \mu(t) = \mathcal{L}(Z(t))$ is continuous and Lipschitz in the Wasserstein metric.

Define, for every real x,

$$F(x) = \mathbb{P}(Z(t) \leq x) \tag{1.28}$$

for any $t > 0$ such that $x + \log_{1/\alpha} t$ is an integer; since $Z(t)$ is periodic in $\log_{1/\alpha} t$, this does not depend on the choice of t.

Since $\widetilde{X}_{\alpha^{-j}t} + \log_{1/\alpha} t = X_{\alpha^{-j}t} - j \in \mathbb{Z}$, the random variable $Z(t) + \log_{1/\alpha} t$ is integer-valued for every t. It is easily seen that for integer-valued random variables Z_1 and Z_2, the total variation distance $d_{\mathrm{TV}}(Z_1, Z_2) \leq d_{\mathrm{W}}(Z_1, Z_2)$. Hence, for any $x \in \mathbb{R}$ and $u \geq 0$, choosing t such that $x + \log_{1/\alpha} t$ is an integer and letting $\gamma = \alpha^{-u}$, which implies that $x - u + \log_{1/\alpha}(\gamma t) = x + \log_{1/\alpha}(t) \in \mathbb{Z}$, we obtain from the definition (1.28) and (1.27),

$$\begin{aligned} F(x) - F(x-u) &= \mathbb{P}\big(Z(t) \leq x\big) - \mathbb{P}\big(Z(\gamma t) \leq x - u\big) \\ &= \mathbb{P}\big(Z(t) + \log_{1/\alpha} t \leq x + \log_{1/\alpha} t\big) - \mathbb{P}\big(Z(\gamma t) + \log_{1/\alpha}(\gamma t) \leq x + \log_{1/\alpha} t\big) \\ &\leq d_{\mathrm{TV}}\big(Z(t) + \log_{1/\alpha} t,\ Z(\gamma t) + \log_{1/\alpha}(\gamma t)\big) \\ &\leq d_{\mathrm{W}}\big(Z(t) + \log_{1/\alpha} t,\ Z(\gamma t) + \log_{1/\alpha}(\gamma t)\big) \\ &\leq d_{\mathrm{W}}\big(Z(t), Z(\gamma t)\big) + |\log_{1/\alpha} t - \log_{1/\alpha}(\gamma t)| \\ &= \mathcal{O}(\gamma - 1) + \log_{1/\alpha} \gamma = \mathcal{O}(u). \end{aligned} \tag{1.29}$$

Hence, $F(x)$ is a continuous function of x.

We have shown that $t \to \mathcal{L}(Z(t))$ is continuous in the Wasserstein metric, and thus in the usual topology of weak convergence in the space $\mathcal{P}(\mathbb{R})$ of probability measures on $\mathbb{R}$. Since further $\mathcal{L}(Z(t))$ is periodic in t, the set $\{\mathcal{L}(Z(t)) : t > 0\} = \{\mathcal{L}(Z(t)) : 1 \le t \le \alpha^{-1}\}$ is compact in $\mathcal{P}(\mathbb{R})$, which by Prohorov's theorem means that the family $\{Z(t)\}$ of random variables is tight, see e.g. Billingsley [1]. Hence, $\mathbb{P}(Z(t) \le x) \to 0$ as $x \to -\infty$ and $\mathbb{P}(Z(t) \le x) \to 1$ as $x \to +\infty$, uniformly in t, and it follows from (1.28) that $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

Furthermore, (1.26) and (1.28) show that, for any sequence k_n of integers, as $n \to \infty$,

$$\begin{aligned}\mathbb{P}(X_n \le k_n) &= \mathbb{P}\big(\widetilde{X}_n \le k_n - \log_{1/\alpha} n\big) = \mathbb{P}\big(Z(n) \le k_n - \log_{1/\alpha} n\big) + o(1)\\ &= F(k_n - \log_{1/\alpha} n) + o(1).\end{aligned} \tag{1.30}$$

Since further the sequence X_n is increasing, it now follows from Janson [12, Lemma 4.6] that F is monotone, and thus a distribution function. By (1.29), the distribution is absolutely continuous and has a bounded density function $F'(x)$.

It is easy to see that (1.30), (1.12) and (1.13) are equivalent, see [12, Lemma 4.1]. The corresponding result in the Wasserstein distance follows from (1.26) because $d_{\mathrm{W}}(X_n, \lceil Z+\log_{1/\alpha} n\rceil) = d_{\mathrm{W}}(\widetilde{X}_n, Z(n))$, e.g. by Remark 1.4 below; (1.14) then follows by (1.10). Finally, (1.26) implies that

$$\mathbb{E}\widetilde{X}_t = \mathbb{E}X_{\lfloor t\rfloor} - \log_{1/\alpha} t = \mathbb{E}Z(t) + \mathcal{O}\left(\log^{-\varepsilon} t\right),$$

which proves (1.16) with $\phi(t) := \mathbb{E}Z(t)$, which is periodic in $\log_{1/\alpha} t$. Since $|\phi(t) - \phi(u)| = |\mathbb{E}(Z(t) - Z(u))| \le d_{\mathrm{W}}(Z(t), Z(u))$, (1.27) implies that ϕ is continuous, and Lipschitz on compact intervals. □

Remark 1.4. As remarked above, $Z(t) + \log_{1/\alpha} t$ is integer-valued. Moreover, for every integer k,

$$\begin{aligned}\mathbb{P}\big(Z(t) + \log_{1/\alpha} t \le k\big) &= \mathbb{P}\big(Z(t) \le k - \log_{1/\alpha} t\big) = F(k - \log_{1/\alpha} t)\\ &= \mathbb{P}\big(Z \le k - \log_{1/\alpha} t\big) = \mathbb{P}\big(Z + \log_{1/\alpha} t \le k\big)\\ &= \mathbb{P}\big(\lceil Z + \log_{1/\alpha} t\rceil \le k\big).\end{aligned}$$

Hence, for every $t > 0$, $Z(t) \overset{\mathrm{d}}{=} \lceil Z + \log_{1/\alpha} t\rceil - \log_{1/\alpha} t$. General families of random variables of this type are studied in [12]. In particular, [12, Theorem 2.3] shows how $\phi(t) := \mathbb{E}Z(t)$ in Theorem 1.3 can be obtained from the characteristic function of the distribution F of Z.

Remark 1.5. The very slow convergence rate $\mathcal{O}\left(\log^{-\varepsilon} t\right)$ in (1.26) is because we allow δ_n to tend to 0 slowly. In typical applications, $\delta_n = n^{-a}$ for some $a > 0$, and then better convergence rates can be obtained. We have, however, not pursued this.

Remark 1.6. Note that F and ϕ are influenced by the distribution of Y_n for small $n > 2$, for example Y_3 and Y_4; hence there is no hope for a nice explicit formula for F or ϕ depending only on asymptotic properties of Y_n.

Remark 1.7. The general problem of studying the number of steps until absorption at 1 of a decreasing Markov chain on $\{1, 2, \dots\}$ appears in many other situations too, usually with quite different behaviour of Y_n and X_n. As examples we mention the recent papers studying random trees and coalescents by Drmota, Iksanov, Moehle and Roesler [3], Iksanov and Möhle [10], Gnedin and Yakubovich [8], and Delmas, Dhersin and Siri-Jegousse [11]; in these papers the number killed in each round is much smaller than here and thus X_n is larger, of the order n or $n/\log n$; moreover, after normalization X_n has a stable limit law.

2 Extensions

In Section 1, we have assumed that we repeat the elimination step until only one player remains. As a generalization we may suppose, in the present Section 2, that we stop when there are at most a players left, for some given number a.

Theorem 2.1. *Consider the leader election algorithm described in Section 1, but stopping as soon as the number of remaining players is at most a, for some fixed $a \geq 1$. Suppose that Condition 1.1 is satisfied. Then, the conclusions of Theorem 1.3 hold, for some F and ϕ that depend on the threshold a.*

Proof: This generalization can be obtained from the version in Section 1 by replacing Y_n by

$$Y'_n := \begin{cases} Y_n, & Y_n > a; \\ 1, & Y_n \leq a. \end{cases}$$

Suppose that Condition 1.1 holds for (Y_n). It is easily seen that then Condition 1.1 holds for (Y'_n) too, with the same α; for (ii), note that Condition 1.1(iii) implies that

$$\mathbb{E}|Y_n - Y'_n| \leq a\mathbb{P}(Y_n \leq a) = \mathcal{O}(n^{-2-\varepsilon})$$

and thus $\mathbb{E}Y'_{n+1} - \mathbb{E}Y'_n = \mathbb{E}Y_{n+1} - \mathbb{E}Y_n + o(n^{-2})$. Consequently, Theorem 1.3 applies to (Y'_n), and the result follows. □

In this situation, it is also interesting to study the probability $\pi_i(n)$ that the procedure ends with exactly i players, starting with n players; here $i = 1, \dots, a$ and $\sum_{i=1}^{a} \pi_i(n) = 1$. We have a corresponding limit theorem for $\pi_i(n)$.

Theorem 2.2. *Suppose that Condition 1.1 holds and that $a \geq 1$ is given as in Theorem 2.1. Then,*

$$\pi_i(n) = \psi_i(n) + o(1), \qquad i = 1, \dots, a, \tag{2.1}$$

for some continuous functions $\psi_i(t)$ on $(0, \infty)$ which are periodic in $\log_{1/\alpha} t$, i.e. $\psi_i(t) = \psi_i(\alpha t)$, and locally Lipschitz.

Proof: A modification of the proof of Theorem 1.3, now taking $x_n := \pi_i(n)$ and $d_n := |x_{n+1} - x_n|$ and replacing the random $\widetilde{X}_t$ by $x_{\lfloor t \rfloor} = \pi_i(\lfloor t \rfloor)$, yields $\pi_i(n+1) - \pi_i(n) = \mathcal{O}(1/n)$ and $\pi_i(\alpha^{-j}t) - \pi_i(\alpha^{-(j+1)}t) = \mathcal{O}(j^{-1-\varepsilon})$; hence, for any $t > 0$, $\pi_i(\alpha^{-j}t) \to \psi_i(t)$ for some $\psi_i(t)$, which easily is seen to satisfy the stated conditions. We omit the details. □

More generally, there is a similar result on the probability that the process passes through a certain state; this is interesting also for the process in Section 1 with $a = 1$.

Theorem 2.3. *Suppose that Condition 1.1 holds and that $a \geq 1$ is given as above. Let $\pi_i(n)$, $i \geq 1$, be the probability that, starting with n players, there exists some round with exactly i survivors. Then*

$$\pi_i(n) = \psi_i(n) + o(1), \qquad i = 1, 2, \dots, \tag{2.2}$$

for some continuous functions $\psi_i(t)$ on $(0, \infty)$ which are periodic in $\log_{1/\alpha} t$, i.e. $\psi_i(t) = \psi_i(\alpha t)$, and locally Lipschitz.

Proof: For $i \le a$, this $\pi_i(n)$ is the same as in Theorem 2.2, and for each $i > a$, this $\pi_i(n)$ is the same as in Theorem 2.2 if we replace a by i. $\square$

Remark 2.4. Another variation, which is natural in some problems, is to study a non-increasing Markov chain on $\{0, 1, \dots\}$ and ask for the number of steps to reach 0; in other words, the time until all players are killed. In this case, we thus assume that $0 \le Y_n \le n$. This can obviously be transformed to our set-up on $\{1, 2, \dots\}$ by increasing each integer by 1; in other words, we replace Y_n by $Y'_n := Y_{n-1} + 1$, $n \ge 2$; we can interpret this as adding a dummy player that never is eliminated, and continuing until only the dummy remains. If Condition 1.1 holds for Y_n, except that $Y_n = 0$ is allowed and $\mathbb{P}(Y_1 = 0) > 0$, then Condition 1.1 holds for Y'_n too, and thus our results hold also in this case, with X_n now defined as the number of steps until absorption in 0. (To be precise, $X_n = X'_{n+1}$, with (X'_n) corresponding to (Y_n), since we add a dummy, but there is no difference between the asymptotics of X'_{n+1} and X'_n.)

3 Examples

Example 3.1 (a toy example)**.** For a simple example to illustrate the theorems above, let, for $n \ge 2$, $Y_n = \lfloor (n+I)/2 \rfloor$, where $I \sim \mathrm{Be}(1/2)$ is 0 or 1 with $\mathbb{P}(I = 1) = 1/2$. In other words, we toss a coin and let Y_n be either $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$ depending on the outcome. (If n is even, thus always $Y_n = n/2$.) Note that $\mathbb{E}Y_n = n/2$, $n \ge 2$, and that Condition 1.1 holds trivially, with $\alpha = 1/2$. If we start with $N_0 = n$ players and $m2^j \le n \le (m+1)2^j$, $m \ge 1$, then the number N_j of survivors after j rounds satisfies $m \le N_j \le m+1$ and $\mathbb{E}N_j = 2^{-j}n$ (by induction on j). Consequently, if $m2^j \le n \le (m+1)2^j$,

$$\mathbb{P}(N_j = m) = m + 1 - 2^{-j}n, \qquad \mathbb{P}(N_j = m+1) = 2^{-j}n - m. \tag{3.1}$$

Taking $m = 1$, this shows that if $2^j \le n \le 2^{j+1}$, then

$$\mathbb{P}(X_n = j) = \mathbb{P}(N_j = 1) = 2 - 2^{-j}n, \quad \mathbb{P}(X_n = j+1) = 1 - \mathbb{P}(X_n = j) = 2^{-j}n - 1. \tag{3.2}$$

Hence, (1.12) holds exactly, $\mathbb{P}(X_n \le k) = F(k - \log_2 n)$, for all $k \in \mathbb{Z}$ and $n \ge 1$, with

$$F(x) = \begin{cases} 0, & x \le -1, \\ 2 - 2^{-x}, & -1 \le x \le 0, \\ 1, & x \ge 0, \end{cases}$$

and (1.16) holds exactly, $\mathbb{E}X_n = \log_2 n + \phi(n)$, with

$$\phi(2^x) = 2^{x - \lfloor x \rfloor} - \big(x - \lfloor x \rfloor\big) - 1.$$

Suppose now, as in Section 2, that we stop when there are at most $a = 3$ players left. (Similar results are easily obtained for other values of a.) If $2^i \le n \le 2^{i+1}$ with $i \ge 1$, we take $j = i - 1$ and note that $N_j \in \{2, 3, 4\}$. If $N_j = 4$, the procedure ends after one further round with 2 players left; otherwise it ends immediately with $N_j = 2$ or 3 survivors. Taking $m = 2$ or $m = 3$ in (3.1), we thus find

$$\pi_2(n) = \mathbb{P}(N_{i-1} = 2) + \mathbb{P}(N_{i-1} = 4) = |2^{1-i}n - 3|, \qquad 2^i \le n \le 2^{i+1}.$$

Consequently, $\pi_2(n) = \psi_2(n)$ exactly, for all $n \ge 2$, with

$$\psi_2(2^x) = |2^{1+x-\lfloor x \rfloor} - 3|.$$

Further, $\psi_3(t) = 1 - \psi_2(t)$ and $\psi_1(t) = 0$.

Example 3.2 (a counter example)**.** The procedure in Example 3.1 is almost deterministic. In contrast, the very similar but completely deterministic $Y_n = \lfloor n/2 \rfloor$, $n \geq 2$, does not satisfy Condition 1.1(ii). In this case, $X_n = \lfloor \log_2 n \rfloor$ and $\mathbb{P}(X_n \leq k) = F(k - \log_2 n)$, for all $k \in \mathbb{Z}$ and $n \geq 1$, where $F(x) = [\![x \geq 0]\!]$, the distribution of $Z := 0$; this limit F is not continuous so the conclusions of Theorem 1.3 do not all hold.

Example 3.3. A leader election algorithm studied by Prodinger [20], Fill, Mahmoud and Szpankowski [5], Knessl [14], and Louchard and Prodinger [18], see also Szpankowski [21, Section 10.5.1], is the following: *Each player tosses a fair coin. If at least one player throws heads, then all players throwing tails are eliminated; if all players throw tails, then all survive until the next round.*

Except for the special rule when all throw tails, which guarantees that at least one player survives each round, the number Y_n of survivors in a round thus has a binomial distribution $\mathrm{Bi}(n, \frac{1}{2})$. More precisely, if W_n is the number of heads thrown,

$$Y_n = W_n + n[\![W_n = 0]\!] \qquad \text{with } W_n \sim \mathrm{Bi}(n, \tfrac{1}{2}). \tag{3.3}$$

Note that

$$\mathbb{E}|Y_n - n/2|^6 = \mathbb{E}|W_n - n/2|^6 = \mathcal{O}(n^3),$$

so (1.4) holds for $p = 6$ (and, indeed, for any $p > 0$), and thus (1.3) holds. Similarly,

$$\mathbb{E}Y_n = \mathbb{E}W_n + 2^{-n}n = \tfrac{1}{2}n + n2^{-n}.$$

Thus, conditions (ii) and (iii) in Theorem 1.3 are satisfied. Also the monotonicity condition (i) is satisfied, because if $1 \leq k \leq n-1$ (other cases are trivial), then

$$\begin{aligned}\mathbb{P}(Y_{n+1} \leq k) &= \mathbb{P}(1 \leq W_{n+1} \leq k) = \tfrac{1}{2}\mathbb{P}(1 \leq W_n \leq k) + \tfrac{1}{2}\mathbb{P}(0 \leq W_n \leq k-1) \\ &= \mathbb{P}(Y_n \leq k) + \tfrac{1}{2}\big(\mathbb{P}(W_n = 0) - \mathbb{P}(W_n = k)\big) \\ &< \mathbb{P}(Y_n \leq k).\end{aligned} \tag{3.4}$$

Hence Condition 1.1 is satisfied with $\alpha = 1/2$ and Theorem 1.3 applies.

In this case, Prodinger [20], see also Fill, Mahmoud and Szpankowski [5], found an exact formula for the expectation $\mathbb{E}X_n$ and asymptotics of the form (1.16) with the explicit function

$$\phi(t) = \tfrac{1}{2} - (\log 2)^{-1} \sum_{k \neq 0} \zeta(1 - \chi_k)\Gamma(1 - \chi_k)e^{2k\pi \mathbf{i} \log_2 t}, \qquad \chi_k := \frac{2k\pi}{\log 2}\mathbf{i}. \tag{3.5}$$

Fill, Mahmoud and Szpankowski [5] further found asymptotics of the distribution that can be written as (1.12) with

$$F(x) = \frac{2^{-x}}{\exp(2^{-x}) - 1}, \tag{3.6}$$

which thus is the distribution function of Z in this case. Second and higher moments are considered by Louchard and Prodinger [18].

Prodinger [20] considered also the possibility of stopping at $a = 2$ players, and showed (1.16) above in this case too, with an explicit formula for the function $\phi(t)$ (of the same type as (3.5) for $a = 1$).

Remark 3.4. Grabner [9] studied a variation of the above Example 3.3, where a given number b of players is to be selected. (The process runs as in Section 1 as long as there are more than b survivors; if exactly b remain, they are selected and the algorithm terminates; if $j < b$ survivors remain, they are selected and we proceed to select $b - j$ further players among the ones eliminated in the last step.) It seems likely that this algorithm (and similar ones) can be treated by an extension of the arguments in this paper, using induction on b, but we have not studied the details.

Example 3.5. A variation of Example 3.3 studied by Janson and Szpankowski [13], Knessl [14] and Louchard and Prodinger [18] is to let the coin be biased, with probability $p \in (0, 1)$ for heads (=survival). Then (3.3) still holds, but with $W_n \sim \operatorname{Bi}(n, p)$. Conditions 1.1(ii)(iii) hold as above, with $\alpha = p$, but arguing as in (3.4) we see that Condition 1.1(i) holds if $p \geq 1/2$, but not for smaller p. Hence, Theorem 1.3 applies when $p \geq 1/2$.

In fact, the results of [13] show that the conclusions (1.12) and (1.16) hold for all $p \in (0, 1)$, for some functions F and ϕ explicitly given in [13]. (See also [18], where further higher moments are treated.) This suggests that Theorem 1.3 should hold more generally. Note that although Condition 1.1(i) does not hold for $p < 1/2$, the difference $\mathbb{P}(Y_{n+1} \leq k) - \mathbb{P}(Y_n \leq k)$ is at most $\mathbb{P}(Y_n = 0) = (1 - p)^n$, and thus negative or exponentially small for large n. It seems likely that Theorem 1.3 can be extended to such cases, by allowing a small error in Condition 1.1(i); this then would include this leader election algorithm with a biased coin for any $p \in (0, 1)$. However, we have not pursued this.

It was left as an open question in [13] whether for each $p \in (0, 1)$ the limit function F is monotone, and thus a distribution function, which means that there exists a random variable Z such that (1.13) holds. By the discussion above, Theorem 1.3 shows that this holds for $p \geq 1/2$, but the case $p < 1/2$ is as far as we know still open. Cf. Remark 4.3 and Figure 6 below, which show that monotonicity fails in a related situation. Numerical experiments, based on [18, Prop. 3.1], indicate that F is monotone, at least for some choices of $p < 1/2$.

A further variation of Example 3.3 is to let the probability p depend on n. The case $p = 1/n$ is studied by Lavault and Louchard [15]; in this case $\mathbb{E}Y_n$ is bounded and Condition 1.1 does not hold, so Theorem 1.3 does not apply.

Example 3.6. The special rule in Examples 3.3 and 3.5 for the exceptional case when all throw tails is of course necessary to prevent us from killing all players, but as we have seen, it complicates the analysis, especially for $p < 1/2$ when it destroys stochastic monotonicity of Y_n. Note that this rule typically is invoked only towards the end of the algorithm, when only a few players are left. We regard the rule as an emergency exit, and it could be replaced by other special rules for this case. For example, an alternative would be to switch to some other algorithm that is fail-safe although in principle (for large n) slower; for our purpose this means that the present algorithm terminates, so we may describe this by letting $Y_n = 1$ in this case, i.e., (3.3) is replaced by $Y_n = W_n + [\![W_n = 0]\!] = \max(W_n, 1)$. Note that for this version, Condition 1.1 holds for every $p \in (0, 1)$, with $\alpha = p$, so our theorems apply.

An equivalent way to treat this version is to add (as in Remark 2.4) a dummy, which is exempt from elimination, and to eliminate everyone else that throws a tail; we then stop when there are at most 2 players left (the dummy and, possibly, one real player). This is thus the version in Section 2, with $a = 2$ and $Y_n = 1 + W_{n-1}$, $W_m \sim \operatorname{Bi}(m, p)$ (and starting with $n + 1$ players). Again, Condition 1.1 holds, with $\alpha = p$, and the results in Section 2 apply. In particular, since invoking the special rule corresponds to eliminating everyone except the dummy, the probability that we have to invoke the special rule is the same as the probability that the dummy version ends with only the dummy, i.e. $\pi_1(n + 1)$ in the notation of Theorem 2.2; the asymptotics of this probability is thus given by (2.1).

Example 3.7. Prodinger [19] (for $p = 1/2$) and Louchard and Prodinger [17] studied a version of Examples 3.3 and 3.5 where, as in Remark 2.4, we allow all players to be killed and let X_n be the time until that happens. Thus, in each round, each player tosses a coin and is killed with probability $1 - p$ (and we do not have any special rule). Additionally, there is a demon, who in each round kills one of the survivors (if any) with probability $\nu \in [0, 1]$. We thus have the modification discussed in Remark 2.4, with

$$Y_n = \max(W_n - I, 0), \qquad W_n \sim \mathrm{Bi}(n, p),\ I_\nu \sim \mathrm{Be}(\nu),$$

where W_n and I_ν are independent; thus also $Y'_n = Y_{n-1} + 1 = \max(W_{n-1} + 1 - I_\nu, 1)$. Condition 1.1 holds, in the modification for absorption in 0, and thus our results apply. In fact, Louchard and Prodinger [17] show (1.12) and (1.16) for this problem with explicitly given F and ϕ; they further give an extension of (1.16) to higher moments.

As remarked in [19, 17], the special case $\nu = 1$ is equivalent to approximate counting and $\nu = 0$ is equivalent to the cost of an unsuccessful search in a trie; in the latter case, X_n is simply the maximum of n i.i.d geometric random variables which can be treated by elementary methods.

4 An application: Variations of Franklin's leader election algorithm

4.1 The algorithms

Franklin's original algorithm. W.R. Franklin [7] proposed a leader election algorithm where the n players are arranged in a ring. Each player gets a random number; these are i.i.d. and, say, uniform on $[0, 1]$. (Since only the order of these numbers will matter, any continuous distribution will do; moreover, it is equivalent to let $\xi_1, \dots, \xi_n$ be a random permutation of $1, \dots, n$.) A player survives the first round if her random number is a peak; in other words, if $\xi_1, \dots, \xi_n$ are i.i.d. random numbers, then player i survives if $\xi_i \geq \xi_{i-1}$ and $\xi \geq \xi_{i+1}$ (with indices taken modulo n). We may ignore the possibility that two numbers ξ_i and ξ_j are equal; hence we may as well require $\xi_i > \xi_{i-1}$ and $\xi > \xi_{i+1}$.

In Franklin's algorithm, the survivors continue by comparing their original numbers in the same way with the nearest surviving players; this is repeated until a single winner remains. We have so far not been able to analyse this algorithm. It is easy to verify that even if we condition on the number m of survivors after the first round, the $m!$ possible different orderings of the survivors do not appear with equal probabilities, which means that the algorithm is *not* of the recursive type studied in this paper. (For example, starting with a ring of 8 players and conditioning on having 4 survivors (peaks) in the first round; the probability of getting 2 survivors in the second round is 10/34, and not 1/3 as in the uniform case.)

A variation. However, we can study a variation of Franklin's algorithm, where the survivors draw new random numbers in each round. This is an algorithm of the type studied in Section 1, with Y_n given by the number of peaks in a random permutation, regarded as a circular list. Note that there is always at least one peak (the maximum will always do), so $Y_n \geq 1$ as required. It is easily seen that inserting a new player will never decrease the number of peaks; hence Y_n is stochastically increasing in n. Further, we have $Y_n = \sum_{i=1}^n I_i$, where $I_i := [\![\xi_i > \max(\xi_{i-1}, \xi_{i+1})]\!]$ is the indicator that player i survives (again, indices are taken modulo n). If $n \geq 3$, then $\mathbb{E} I_i = 1/3$ by symmetry, and thus $\mathbb{E} Y_n = n/3$. In particular, (ii) holds with $\alpha = 1/3$. Furthermore, I_i and I_j are independent unless $|i - j| \leq 2 \pmod n$, and similarly for sets of the indicators I_i, and it follows easily that $\mathbb{E}(Y_n - \mathbb{E}Y_n)^6 = \mathcal{O}(n^3)$, and thus (1.4) holds for $p = 6$. (Indeed, (1.4) holds for all even p by this argument.) Consequently, Condition 1.1 holds and Theorem 1.3 applies, with $1/\alpha = 3$.

Remark 4.1. It might be shown that for the true Franklin algorithm, the expected number of survivors after 2 rounds is $c_2 n + o(n)$, where

$$c_2 = \frac{3e^4 - 48e^2 + 233}{384} \approx 0.1096868681. \tag{4.1}$$

(Details might appear elsewhere.) In comparison, for the variation with new random numbers each round, it is easily seen that the expected number after k rounds is $(1/3)^k n + o(1)$, for any fixed k; in particular, after two rounds it is $n/9 + o(n)$. Note that c_2 in (4.1) is slightly smaller than $1/9$, and thus better in terms of performance. It might have been hoped that the original Franklin algorithm is asymptotically equivalent to the variation studied here, but the fact that $c_2 \neq 1/9$ suggests that this is not the case. Nevertheless, we conjecture that Theorem 1.3 still holds for the true Franklin algorithm, for some unknown $\alpha < 1/3$.

The linear case. In both versions above of Franklin's algorithm, the players are arranged in a circle. Alternatively, the players may be arranged *in a line*. We use the same rules as above, but we have to specify when a player at the end (with only one neighbour) is a peak. There are two obvious possibilities:

(i) Never regard the first and last players as peaks. (Define $\xi_0 = \xi_{n+1} = +\infty$.)

(ii) Regard them as peaks if $\xi_1 > \xi_2$ and $\xi_n > \xi_{n-1}$, respectively. (Define $\xi_0 = \xi_{n+1} = -\infty$.)

In the first case, it is possible that there are no peaks, and thus we have to add an emergency exit as in Example 3.6.

As in the circular case, there are two versions (for each of (i) and (ii)): we may use the same random numbers in all rounds, or we may draw new ones each round.

In the latter case, we are again in the situation of Section 1. In both cases (i) and (ii), the distribution of Y_n is related to the distribution in the circular case discussed above. Indeed, if we start with a circular list of $n + 1$ numbers and eliminate the player with the largest number, then the remaining n numbers form a linear list, and the peaks in this list using version (i) equal the peaks except the maximum one in the original circular list. Similarly, if we instead eliminate the player with the smallest number, then the peaks in the remaining list using version (ii) equal the peaks in the original list. Hence, if Z_n is the number of peaks in a random circular list of length n, then (i) yields $Y_n \overset{\mathrm{d}}{=} Z_{n+1} - 1$ and (ii) yields $Y_n \overset{\mathrm{d}}{=} Z_{n+1}$. In both cases, this implies that Condition 1.1 holds (provided we add a suitable emergency exit in case (i)) because it holds for Z_n. Consequently, Theorem 1.3 applies to both these linear versions of (the variation of) Franklin's algorithm, again with $1/\alpha = 3$.

In the following subsections we analyze further two of these variations of Franklin's algorithm (both with new random numbers drawn in each round), using numerical calculations.

4.2 First variation of Franklin's algorithm: the linear case

We assume that the survivors draw new random numbers in each round and that they are arranged in line. We use possibility (i) in Section 4.1. We start with a set of n players. We assign a classical permutation of $\{1, \ldots, n\}$ to the set, all players corresponding to a peak stay alive, the other ones are killed. If there are no peaks, we choose the following emergency exit: a player is chosen at random (this is assumed to have 0 cost), indeed in the original game, one deals with circular permutations, so there always exists at least one peak, here we approach the problem with a classical inline permutation.

What is the distribution of the number X_n of *phases* (or rounds) before getting only one player?

4.2.1 The analysis

Let

$$\begin{aligned} Y_n &:= \text{ number of peaks, starting with } n \text{ players},\\ P(n,k) &:= \mathbb{P}[Y_n = k] = \mathbb{P}[k \text{ peaks, starting with } n \text{ players}],\\ \Pi(n,j) &:= \mathbb{P}(X_n = j) = \mathbb{P}[j \text{ phases are necessary to end the game, starting with } n \text{ players}],\\ \Lambda(n,j) &:= \sum_{k=0}^{j} \Pi(n,k) = \mathbb{P}[\text{at most } j \text{ phases are necessary, starting with } n \text{ players}]. \end{aligned}$$

We will sometimes use the subscript ℓ to distinguish these from the circular case discussed in Section 4.3.

First of all, we know (Carlitz [2]; see also [6, Chapter 3]), that the pentavariate generating function (GF) of valleys (u_0), double rise (u_1), double fall (u_1') and peaks (u_2) is given by

$$I(z,\mathbf{u}) = \frac{\delta}{u_2}\frac{v_1 + \delta\tan(z\delta)}{\delta - v_1\tan(z\delta)} - \frac{v_1}{u_2},$$

with

$$v_1 = (u_1 + u_1')/2, \quad \delta = \sqrt{u_0 u_2 - v_1^2}.$$

This gives the GF of the number of peaks:

$$\frac{\tan[z(u-1)^{1/2}]}{(u-1)^{1/2} - \tan[z(u-1)^{1/2}]}, \tag{4.2}$$

hence the mean $\mathbb{M}$ and variance $\mathbb{V}$ of the number of peaks, for $n \geq 2$ and $n \geq 4$, respectively:

$$\mathbb{M}(n) = (n-2)/3, \quad \mathbb{V}(n) = 2(n+1)/45. \tag{4.3}$$

This GF is also given in Carlitz [2]. Moreover, from [6, Chapter 9], we know that the distribution P is asymptotically Gaussian. This is also proved in Esseen [4] by probabilistic methods.

Let $x(n)$ be the mean number of phases, $\mathbb{E}(X_n)$, starting with n players. As we shall see, the initial values are

$$x(0) = x(1) = 0, \quad x(2) = x(3) = x(4) = 1.$$

Since (4.3) yields $M(n) + 1 = (n+1)/3$ for $n \geq 2$, we have (approximating by using this for $n \leq 1$ too) that the mean number of players $c(j)$ still alive after j phases is

$$c(j) \approx 3^{-j}(n+1) - 1.$$

(An induction easily yields the exact formula $|c(j) - 3^{-j}n| < 1$ for all j.) If we want $c(j) = 1$, this leads to the approximation

$$x(n) \approx j \approx \log_3 n - \log_3 2.$$

$n \backslash k$	0	1	2	3	4
1	1	0	0	0	0
2	1	0	0	0	0
3	$2/3$	$1/3$	0	0	0
4	$1/3$	$2/3$	0	0	0
5	$2/15$	$11/15$	$2/15$	0	0
6	$2/45$	$26/45$	$17/45$	0	0
7	$4/315$	$38/105$	$4/7$	$17/315$	0

Tab. 1: $P_\ell(n,k)$

$n \backslash j$	0	1	2	3
0	1	0	0	0
1	1	0	0	0
2	0	1	0	0
3	0	1	0	0
4	0	1	0	0
5	0	$13/15$	$2/15$	0
6	0	$28/45$	$17/45$	0
7	0	$118/315$	$197/315$	0
$\vdots$				
20	0	$< 10^{-7}$	$\bullet$	$\bullet$

Tab. 2: $\Pi_\ell(n,j)$

We see from Theorem 1.3 (which applies by Section 4.1) that this is roughly correct, but the constant $-\log_3 2$ has to be replaced by a periodic function $\phi(n)$. Let us now construct Π. We have, for $n \geq 2$ and $j \geq 1$,

$$\Pi(n,j) = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} P(n,k)\Pi(k,j-1). \tag{4.4}$$

We have the initial values

$$\Pi(0,0) = 1,\ \Pi(0,j) = 0, j > 0,\ \Pi(1,0) = 1,\ \Pi(1,j) = 0, j > 0,\ \Pi(n,0) = 0, n \geq 2.$$

Also

$$\Pi(2,1) = \Pi(3,1) = \Pi(4,1) = 1.$$

Some values of P and Π are given in Tables 1 and 2. (We use in the tables and figures the subscript ℓ to emphasise that we deal with the linear case.)

Denoting the jth column of Π by $\pi^{(j)}$, we have

$$\pi^{(j)} = P^{j-1}\pi^{(1)}.$$

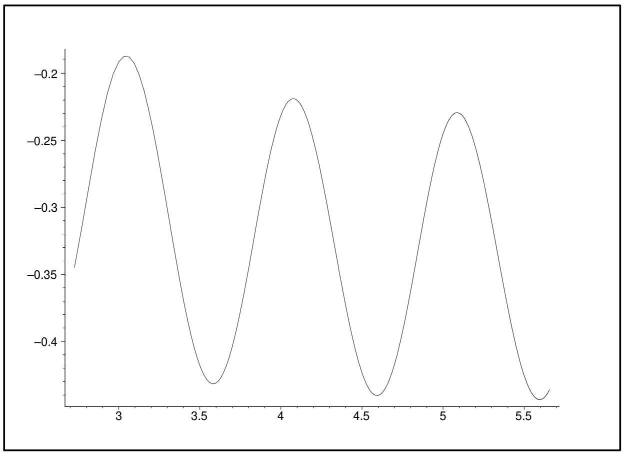


Fig. 2: $x_\ell(n) - \log_3 n$ versus $\log_3 n$, $n = 50, \dots, 500$

For $j \geq 2$, it suffices to consider $k \geq 2$ in (4.4), so we need only the matrix $(P(n,k))_{n,k\geq 2}$. Since $P(n,k) = 0$ if $k > (n-1)/2$, this matrix is triangular, and so is P^{j-1}.

But $\pi^{(1)}(n) < 10^{-7}$, $n > 20$, so numerically, the significant columns of P^{j-1} are the first 20 columns. Also, we see the importance of the initial first column of Π. Moreover, for $n > 75$, $P(n,k)$ is indistinguishable from the Gaussian limit. So we have used the expansion of the GF (4.2) for $n \leq 75$ and the Gaussian limit afterwards in our numerical calculations. Of course we have

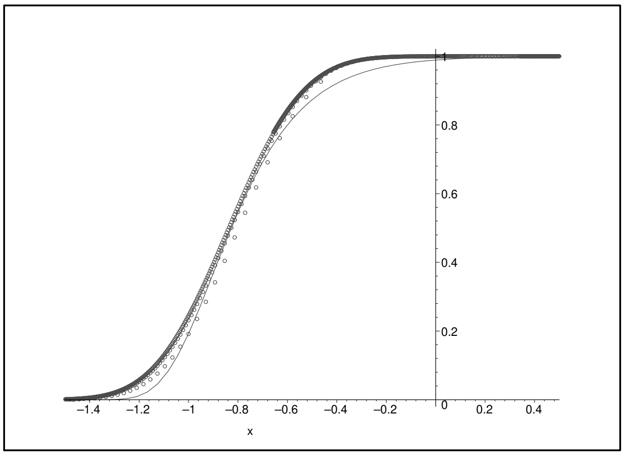
$$x(n) = \sum_{j=0}^{\infty} \Pi(n,j) j,$$

and

$$x(n) = 1 + \sum_{k=2}^{\infty} P(n,k) x(k), \qquad n \geq 2.$$

Remark 4.2. Another approach could be the following: Let

$$\begin{aligned}
I(k) &:= [\![\text{one of the phases has } k \text{ players}]\!], \\
I(j,k) &:= [\![\text{phase } j \text{ has } k \text{ players}]\!], \\
Q(k) &:= \mathbb{P}[I(k) = 1], \\
R(j,k) &:= \mathbb{P}[I(j,k) = 1], \\
Q(k) &= \mathbb{P}\Big[\bigvee_j I(j,k) = 1\Big] = \sum_{j=1}^{n-k} R(j,k), \text{ as } \mathbb{P}[I(j,k) \wedge I(i,k)] = 0, \text{ if } i \neq j, \\
R(j,k) &= \sum_l R(j-1,l) P(l,k), \; j \geq 1, \text{ and } R(0,k) = [\![k = n]\!], \\
x(n) &= \mathbb{E}\Big[\sum_{k=1}^{n} I(k)\Big] = \sum_{k=1}^{n} Q(k).
\end{aligned}$$



$\circ$: observed
— : Gumbel distribution

Fig. 3: $\Lambda_\ell(n,j) = \mathbb{P}(X_n \le j)$ versus $j - \log_3 n$, approximating $F_\ell(x)$, $n = 20, \dots, 500$

A plot of $x(n) - \log_3 n$ versus $\log_3 n$ is given in Figure 2 for $n = 50, \dots, 500$. The oscillations expected from (1.16) are clear.

Recall that according to Theorem 1.3, there exists a limiting distribution function $F(x) = F_\ell(x)$ (in a certain sense) for X_n. In Figure 3, we approximate this distribution function $F(x)$ by plotting $\Lambda(n,j) = \mathbb{P}(X_n \le j)$ against $j - \log_3 n$ for $n = 20, \dots, 500$, cf. (1.12). We have also plotted a scaled Gumbel distribution; the fit is bad.

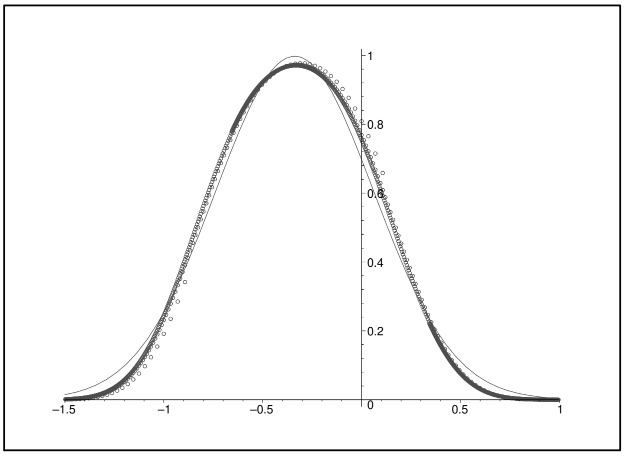
Similarly, in Figure 4 we show the probability $\Pi(n,j)$, $n = 150, \dots, 500$, plotted against $j - \log_3 n$. The fit with a Gaussian distribution is equally bad.

The few scattered points of both figures are actually due to small n and the propagation of the more erratic behaviour for $n = 1, \dots, 40$ shown in Figure 5.

So we observe the following facts:

$n \backslash j$	0	1	2	3
0	0	1	0	0
1	1	0	0	0
2	0	0	1	0
3	0	$1/3$	$2/3$	0
4	0	$2/3$	$1/3$	0
5	0	$11/15$	$2/15$	$2/15$
$\vdots$				

Tab. 3: $\Pi(n,j)$, other initialization



∘ : observed
— : Gaussian distribution

Fig. 4: $\Pi_\ell(n,j) = \mathbb{P}(X_n = j)$ versus $j - \log_3 n$, approximating $\Delta F_\ell(x)$, $n = 150, \ldots, 500$

(i) A first regime ($n = 1, \ldots, 40$) creates some scattered points which almost look like two distributions.

(ii) Between $n = 40$ and $n \sim 75$, a limiting distribution is attained; $\Pi(n,j)$ is concentrated on $j = \log_3 n + \mathcal{O}(1)$.

(iii) For $n > 75$, the limiting Gaussian for P with its narrow ($\sqrt{n}$) dispersion, intuitively induces, with (4.4), a propagation, with some smoothing, of the previous distribution. We attain the limiting distribution $F(x)$ given by Theorem 1.3.

(iv) At most two values carry the main part of the probability mass $\Pi(n,j)$. This is clear from the observed range of $F(x)$ in Figures 3 and 4.

(v) P is triangular, and so is P^j. Also, $P^j(n,k) = \Theta(1)$, with $k = \mathcal{O}(1)$, only if $j = \log_3 n + \mathcal{O}(1)$.

(vi) As $F(x)$ is absolutely continuous, we can derive, as in [16], modulo some uniform integrability conditions, all (periodic) moments of X_n, in particular $x(n)$.

(vii) The effect of *initial values* is now clear. To illustrate this, we have changed to $\Pi(0,1) = \Pi(1,0) = 1$, which means that we add a cost 1 for the extra selection required when the algorithm terminates with no element left. This leads to Table 3. The equivalent of Figures 3, 4 and 5 is given in Figures 6, 7 and 8. Note that X_n no longer is stochastically monotone in n; we have by definition $X_0 = 1 > 0 = X_1$, and Table 3 shows other examples of non-monotonicity for small n. Moreover, Figure 6 shows that the non-monotonicity persists for large n; we clearly have convergence to a limit function, $G(x)$ say, but the limit is not monotone and thus not a distribution function as in Figure 4 and, more generally, in Theorem 1.3.

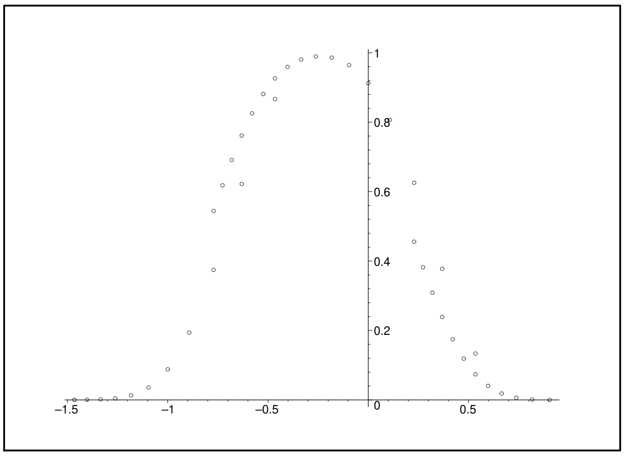


Fig. 5: $\Pi_\ell(n,j)$ versus $j - \log_3 n$, $n = 1, \ldots, 40$

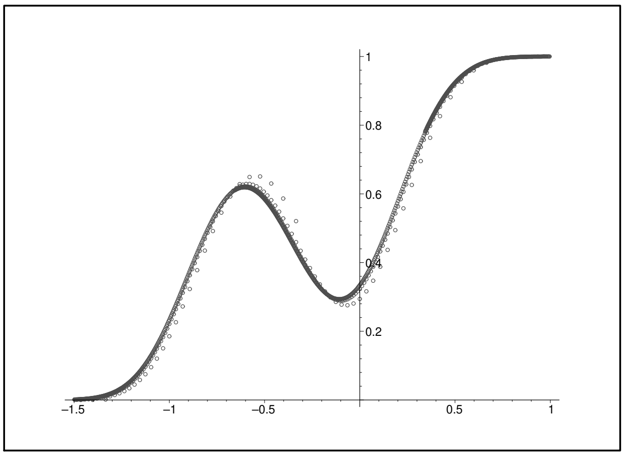


Fig. 6: $\Lambda(n,j)$ versus $j - \log_3 n$, other initialization, $n = 5, \ldots, 500$

Remark 4.3. Note that the example in (vii) and Figure 6 does not contradict Theorem 1.3 because X_n now is defined with other initial values than in Theorem 1.3. Nevertheless, it is a warning that monotonicity of the limit should not be taken for granted in cases such as Example 3.5 with $p < 1/2$ where the monotonicity assumption of Theorem 1.3 is not satisfied.

4.2.2 Periodicities

Let $\psi(\alpha) := \int e^{\alpha x} f(x)\, dx$ be the Laplace transform of the limiting distribution F. (We do not know whether $\psi(\alpha)$ exists in general, although we conjecture so, but we really only need it for imaginary α,

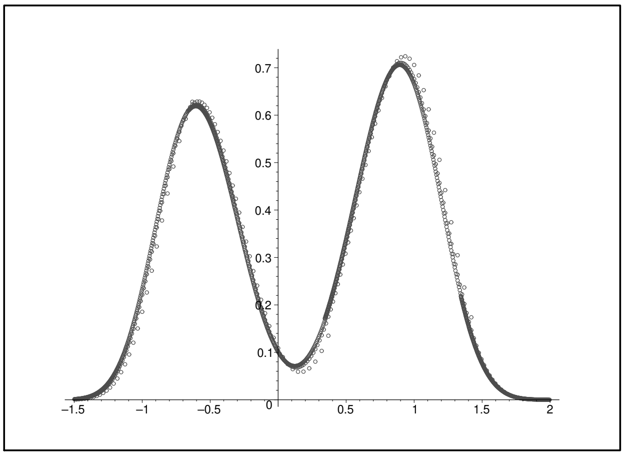


Fig. 7: $\Pi(n, j)$ versus $j - \log_3 n$, other initialization, $n = 5, \ldots, 500$

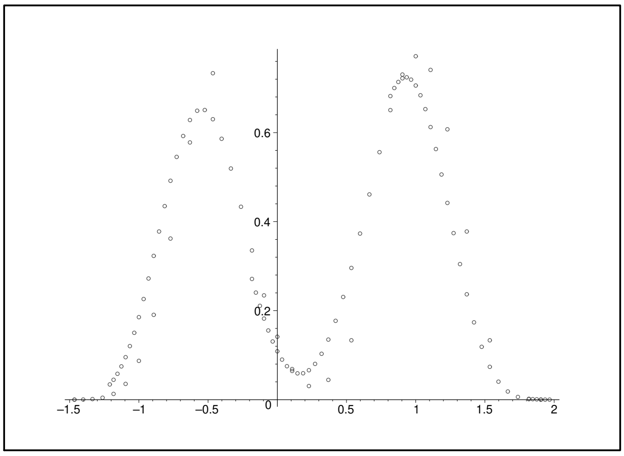


Fig. 8: $\Pi(n, j)$ versus $j - \log_3 n$, $n = 1, \ldots, 100$, other initialization

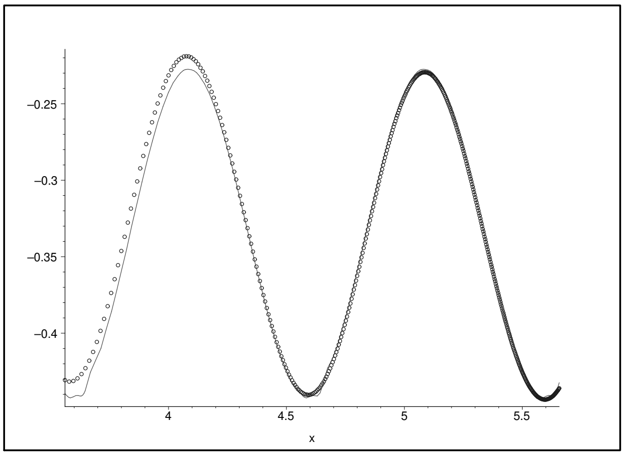


Fig. 9: observed $x_\ell(n) - \log_3 n$ ($\circ$) and computed with (4.5) (line) periodicities versus $\log_3 n$ (linear case), $n = 50, \dots, 500$

i.e., the characteristic function of F.) Similarly, let $\tilde{\psi}(\alpha) := \int e^{\alpha x} \Delta F(x)\, dx$ be the Laplace transform of ΔF. Since $\Delta F(x) := F(x) - F(x-1) = \int_0^1 f(x-t)\, dt$, $\Delta F = f * \mathbf{1}_{[0,1]}$, and thus, since the indicator function $\mathbf{1}_{[0,1]}$ has Laplace transform $(e^\alpha - 1)/\alpha$,

$$\tilde{\psi}(\alpha) = \frac{e^\alpha - 1}{\alpha} \psi(\alpha).$$

With the usual machinery (see [16] or [12, Theorem 2.3]), we obtain from Theorem 1.3 and Remark 1.4, assuming some technical conditions that are very likely to hold but not rigorously verified,

$$x(n) := \mathbb{E} X_n = \log_3 n + \tilde{m}_1 + w_1(\log_3 n) + o(1), \tag{4.5}$$

where

$$\tilde{m}_1 := \tilde{\psi}'(0) = \psi'(0) + \tfrac{1}{2}, \tag{4.6}$$

$$w_1(x) := \sum_{l \neq 0} \tilde{\psi}'(2\pi l \mathbf{i}) e^{2\pi l \mathbf{i} x} = \sum_{l \neq 0} \frac{\psi(2\pi l \mathbf{i})}{2\pi l \mathbf{i}} e^{2\pi l \mathbf{i} x}. \tag{4.7}$$

With the observed values of $\Pi(n,j)$ (see Figure 4), we have computed the Laplace transform numerically (with a modified Euler–MacLaurin, see below) as follows. Assume that we have N computed values $\{\Pi(n_i, j_i) : i = 1, \dots, N\}$. Setting $x = j - \log n$, this gives $\{\Pi(n_i, x_i) : i = 1, \dots, N\}$. Sorting w.r.t. x_i, we write this as $\{\Pi(n_k, y_k) : y_k \le y_{k+1},\ k = 1, \dots, N\}$. Construct a numerical Laplace transform

$$\tilde{\psi}(\alpha) = \sum_k e^{\alpha y_k} \Pi(n_k, y_k)(y_{k+1} - y_{k-1})/2.$$

Using this numerically computed ψ in (4.5)–(4.7), we compute $\tilde{m}_1 + w_1(x)$, which fits quite well with the observed periodicities of $x(n) - \log_3 n$ in Figure 2; the comparison is given in Figure 9.

$n \backslash k$	0	1	2	3	4
1	0	1	0	0	0
2	0	1	0	0	0
3	0	1	0	0	0
4	0	$2/3$	$1/3$	0	0
5	0	$1/3$	$2/3$	0	0
6	0	$2/15$	$11/15$	$2/15$	0
7	0	$2/45$	$26/45$	$17/45$	0

Tab. 4: $P_{\mathrm{c}}(n,k)$

$n \backslash j$	0	1	2	3
0	1	0	0	0
1	1	0	0	0
2	0	1	0	0
3	0	1	0	0
4	0	$2/3$	$1/3$	0
5	0	$1/3$	$2/3$	0
6	0	$2/15$	$13/15$	0
7	0	$2/45$	$43/45$	0
$\vdots$				
20	0	$< 10^{-7}$	$\bullet$	$\bullet$

Tab. 5: $\Pi_{\mathrm{c}}(n,j)$

4.3 Second variation of the Franklin's algorithm: the circular case

If we denote by $P_{\mathrm{c}}(n,k)$ the distribution of the number of peaks in the circular case and by $P_\ell(n,k)$ the distribution in the linear case, we know, by Section 4.1, that $P_{\mathrm{c}}(n,k) = P_\ell(n-1,k-1)$. It is easy to check that this leads to, for $n \geq 3$ and $n \geq 5$, respectively,

$$\mathbb{M}(n) = n/3, \quad \mathbb{V}(n) = 2n/45,$$

which also is easy to see probabilistically, by writing Y_n as the sum of the n indicators $[\![$player i is a peak$]\!]$, and noting that indicators with distance at least 3 are independent.)

The initial values are now given by

$$P(1,1) = 1,\ P(2,1) = 1,\ P(3,1) = 1,\ \Pi(0,0) = 1,\ \Pi(1,0) = 1,\ \Pi(2,1) = 1,\ \Pi(3,1) = 1.$$

The corresponding pictures are given in Tables 4 and 5. (We use a subscript c for the circular case.)

A plot of $\Lambda_{\mathrm{c}}(n,j)$ versus $j - \log_3 n$ for $n = 20, \dots, 500$ is given in Figure 10. We see that there are fewer scattered points than in the linear case. Let us mention that the fits with Gumbel or Gaussian are equally bad. A comparison of $\Lambda_{\mathrm{c}}(n,j)$ with $\Lambda_\ell(n,j)$ is given in Figure 11. No numerical relation exists between the two distributions.

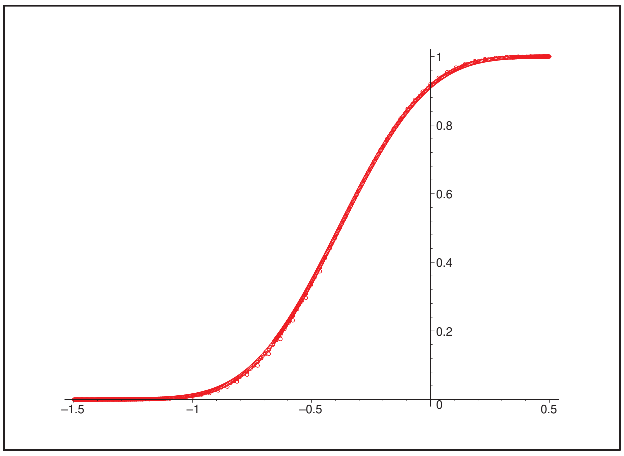


Fig. 10: $\Lambda_c(n, j)$ versus $j - \log_3 n$, approximating $F_c(x)$, $n = 5, \ldots, 500$

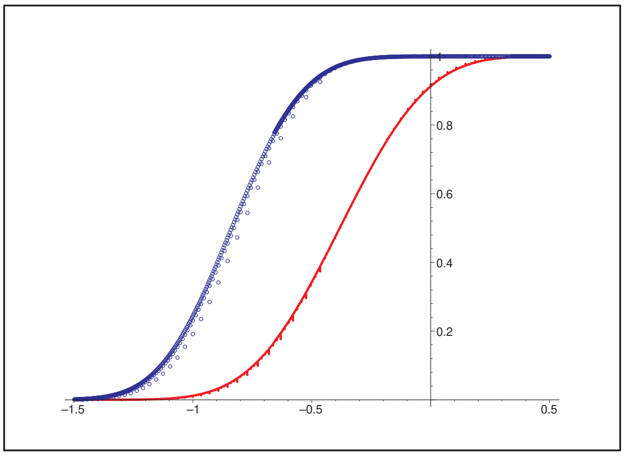


Fig. 11: A comparison of $\Lambda_c(n, j)$ (right curve) with $\Lambda_\ell(n, j)$ (left curve)

$\Pi_c(n, j)$ versus $j - \log_3 n$, $n = 1, \ldots, 40$, is plotted in Figure 12. These initial points are now scattered less than in the linear case.

The observed versus computed periodicities are given in Figure 13.

In conclusion, apart from numerical differences, the behaviour of our two variations are quite similar.

Note that the mean number of needed messages is asymptotically $2n \log_3(n)$, as we use $2n$ messages per round. Franklin [7] gives an upper bound $2n \log_2(n)$.

Acknowledgements

The comments of the referee led to improvements in the presentation.

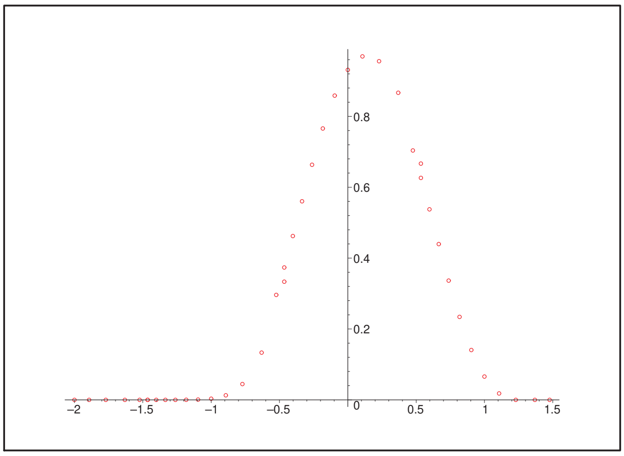


Fig. 12: $\Pi_c(n,j)$ versus $j - \log_3 n$, $n = 1, \ldots, 40$

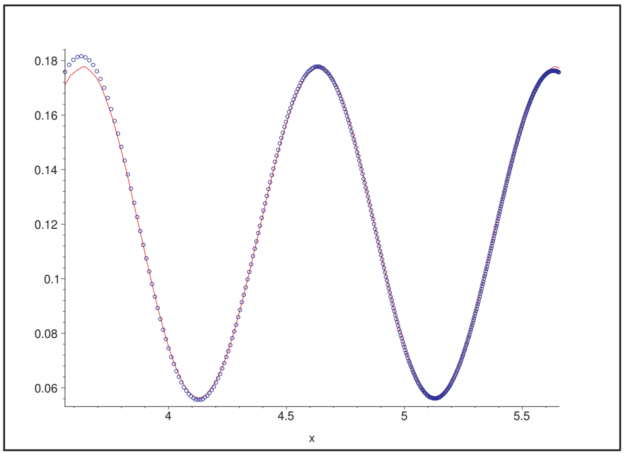


Fig. 13: observed $x_c(n) - \log_3 n$ ($\circ$) and computed with (4.5) (line) periodicities versus $\log_3 n$ (circular case), $n = 50, \ldots, 500$

References

[1] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1968.

[2] L. Carlitz. Permutations and sequences. *Advances in Mathematics*, 14:92–120, 1974.

[3] M. Drmota, A. Iksanov, M. Moehle, and U. Roesler. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures and Algorithms*, to appear.

[4] C. G. Esseen. On the application of the theory of probability to two combinatorial problems invoving permutations. In *Proceedings of the Seventh Conference on Probability Theory (Braşov, 1982)*, pages 137–147, 1982.

[5] J.A. Fill, H. Mahmoud, and W. Szpankowski. On the distribution for the duration of a randomized leader election algorithm. *Annals of Applied Probability*, 1:1260–1283, 1996.

[6] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, to appear, 2008. Available at http://algo.inria.fr/flajolet/Publications/.

[7] W. R. Franklin. On an improved algorithm for decentralized extrema-finding in a circular configuration of processors. *Communications of the ACM*, 25(5):336–337, 1982.

[8] A. Gnedin and Y. Yakubovich. On the number of collisions in Λ-coalescents. *Electron. J. Probab.*, 12(56):1547–1567, 2007. arXiv:0704.3902v1.

[9] P.J. Grabner. Searching for losers. *Random Structures and Algorithms*, 4(1):99–110, 1993.

[10] A. Iksanov and M. Möhle. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electronic Communications in Probability*, 12(4):28–35, 2007.

[11] J.-S. Dhersin J.-F. Delmas and A. Siri-Jegousse. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Annals of Applied Probability*, 18(3):997–1025, 2008.

[12] S. Janson. Rounding of continuous random variables and oscillatory asymptotics. *Annals of Probability*, 34:1807–1826, 2006.

[13] S. Janson and W. Szpankowski. Analysis of an asymmetric leader election algorithm. *Electronic Journal of Combinatorics*, 4(R17), 1997.

[14] C. Knessl. Asymptotics and numerical studies of the leader election algorithm. *European Journal of Applied Mathematics*, 12:645–664, 2001.

[15] C. Lavault and G. Louchard. Asymptotic analysis of a leader election algorithm. *Theoretical Computer Science*, 359:239–254, 2006.

[16] G. Louchard and H. Prodinger. Asymptotics of the moments of extreme-value related distribution functions. *Algorithmica*, 46:431–467, 2006. Long version: http://www.ulb.ac.be/di/mcs/louchard/moml.ps.

[17] G. Louchard and H. Prodinger. Advancing in the presence of a demon. *Mathematica Slovaca*, 58(3):263–276, 2008.

[18] G. Louchard and H. Prodinger. The asymmetric leader election algorithm. *Annals of Combinatorics*, to appear.

[19] H. Prodinger. How to advance on a stairway by coin flippings. In *Applications of Fibonacci Numbers, Vol. 5 (St. Andrews, 1992)*, pages 473–479, 1993.

[20] H. Prodinger. How to select a loser. *Discrete Mathematics*, 120:149–159, 1993.

[21] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences*. Wiley, New York, 2001.