

## **Almost Sure Convergence for Iterated Functions of Independent Random Variables**

Jonathan Jordan  
Basic Research Institute in the Mathematical Sciences  
HPLaboratories Bristol  
HPL-BRIMS-2000-14  
12<sup>th</sup> June, 2000\*

hierarchical  
systems,  
asymptotic  
behaviour,  
laws of large  
numbers

We consider a class of probabilistic models obtained by iterating random functions of  $k$  random variables. We prove an analogue of the weak law of large numbers and under a symmetry condition we prove a strong law. The symmetry condition is satisfied if the initial random variables are exchangeable. Our results can be used to give stronger results than those previously obtained in the special case where the function is deterministic. Both types of model have applications in physics and in computer science.

# ALMOST SURE CONVERGENCE FOR ITERATED FUNCTIONS OF INDEPENDENT RANDOM VARIABLES

JONATHAN JORDAN

ABSTRACT. We consider a class of probabilistic models obtained by iterating random functions of  $k$  random variables. We prove an analogue of the weak law of large numbers and under a symmetry condition we prove a strong law. The symmetry condition is satisfied if the initial random variables are exchangeable. Our results can be used to give stronger results than those previously obtained in the special case where the function is deterministic. Both types of model have applications in physics and in computer science.

## 1. INTRODUCTION

Let  $\mathcal{D} \subseteq (-\infty, \infty)$  be a closed domain and consider a function

$$f : \mathcal{D}^k \rightarrow \mathcal{D} \text{ for some } k \in \mathbb{N}.$$

Now consider a set of i.i.d. random variables  $\{X_j^{(0)}; j \geq 1\} = \mathcal{X}^{(0)}$ , say, on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Define

$$X_j^{(1)} = f(X_{(j-1)k+1}^{(0)}, \dots, X_{jk}^{(0)}) \text{ for } j \geq 1.$$

This gives a sequence  $\mathcal{X}^{(1)} = \mathcal{R}\mathcal{X}^{(0)}$  of i.i.d. random variables. We iterate the map  $\mathcal{R}$  to get a sequence  $\mathcal{X}^{(0)}, \mathcal{X}^{(1)}, \dots, \mathcal{X}^{(n)}, \dots$  of i.i.d. sequences, and we are interested in the limiting behaviour.

---

1991 *Mathematics Subject Classification.* Primary 60F15; secondary 60F05, 60K35, 60K37.

*Key words and phrases.* Hierarchical systems, asymptotic behaviour, laws of large numbers.

Supported by an EPSRC CASE studentship in partnership with Hewlett-Packard's Basic Research Institute in the Mathematical Sciences.

This type of model arises naturally in the study of networks of random resistors on hierarchical lattices, see for example Schlösser and Spohn (1992); Essoh and Bellisard (1989); Schenkel, Wehr and Wittwer (2000); Wehr (1997); Wehr and Woo (2000). Hierarchical models are studied in statistical physics because they can often be solved explicitly and may provide insight into behaviour on standard lattices. For a discussion see Stinchcombe and Watson (1976); Bernasconi (1978) and for studies of some statistical physics models in the hierarchical setting see Griffiths and Kaufman (1982); Derrida (1986). Similar models also occur in other situations arising in physics, such as modelling earthquakes and fibre strengths in Newman et. al. (1994). A discussion of the relationship between classical probability results such as the central limit theorem and renormalization in physics appears in Jona-Lasinio (1975). Fluctuation theorems for specific models are obtained in Essoh and Bellisard (1989); Wehr and Woo (2000); Schenkel, Wehr and Wittwer (2000).

An application to computer science is the study of the biased coin problem considered in Alon and Rabin (1989); Boppana and Narayanan (1993). Furthermore, the model analysed similarly in Alon and Naor (1993); Boppana and Narayanan (1993, 2000) to study collective coin-flipping and leader-election protocols is an example of the randomized hierarchical models also considered in this paper.

In the case where the function is continuous, defined on  $\mathcal{D} = [0, 1]$ , concave, positively homogeneous and increasing, convergence in probability to a constant is proved in Shneiberg (1986). For one particular function obtained by considering a specific lattice, almost sure convergence to a constant is proved in Essoh and Bellisard (1989). The most general result to date is due to Li and Rogers (1999) who at each level  $n$  of the renormalization considered a (deterministic) function  $f_n$  of  $k_n$  variables. They assumed that the initial sequence of random variables was stationary and  $m$ -dependent, i.e. the two collections  $\{X_1^{(0)}, \dots, X_n^{(0)}\}$  and  $\{X_{n+m+1}^{(0)}, X_{n+m+2}^{(0)}, \dots\}$  are independent, and proved a weak law of large numbers. Under stronger conditions they proved the following strong law (Theorem 2.1(iii) of Li and Rogers (1999)).

**Theorem 1.1.** *Suppose that the  $f_n$  satisfy the subadditive constraint*

$$f_n(x_1, \dots, x_k) \leq \frac{1}{k} \sum_{i=1}^k x_i \text{ for all } (x_1, \dots, x_k) \in \mathcal{D}^k.$$

Then if  $\mathcal{D}$  is bounded below, the  $f_n$ ,  $n \geq 1$  are symmetric functions of  $k(\geq 2)$  real variables and  $\mathcal{X}^{(0)} = \{X_j^{(0)}; j \geq 1\}$  is a sequence of i.i.d. random variables such that

$$\mathbb{E}|X_1^{(0)}| < \infty,$$

then for some  $\lambda \in \mathcal{D}$ ,

$$\lim_{n \rightarrow \infty} X_1^{(n)} = \lambda \text{ a.s.}$$

We will develop a randomized version of the problem and extend the weak law from Li and Rogers (1999) to this setting. In the randomized setting the symmetry condition needed for the strong law is less rigid and we can then use this strong law to weaken the conditions of Li and Rogers (1999) in the non-randomized model, giving the following result.

**Theorem 1.2.** (a) *If*

$$f(x_1, \dots, x_n) \leq \sum_{i=1}^k \alpha_i x_i$$

where the  $\alpha_i$  are positive constants with  $\alpha_i < 1$  for each  $i$ ,  $\sum_{i=1}^k \alpha_i = 1$  and

$$\mathbb{E} \left( |X_1^{(0)}| (L(|X_1^{(0)}|))^{\delta} \right) < \infty$$

where

$$\delta > 1 \text{ and } L(x) = \max\{1, \log x\},$$

then there exists  $\lambda \in \mathcal{D}$  with

$$X_1^{(n)} \rightarrow \lambda \text{ a.s.}$$

(b) *Further, if  $\alpha_i = \frac{1}{k}$  for all  $i$ , then the conclusion holds under the weaker condition that  $\mathbb{E}|X_1^{(0)}| < \infty$ .*

Part (b) of this was originally claimed in Wehr (1997).

For recent work on this problem in the Banach space setting, see Hamby and O'Connell (2000).

## 2. REQUIRED LEMMAS

The following results and their proofs are extensions of Proposition 3.1 of Li and Rogers (1999) to the case with random weights.

A set of random variables  $\{X_j : j \geq 1\}$  is said to be stationary and  $m$ -dependent if for all  $n$  the sequence  $\{X_{n+j} : j \geq 1\}$  has the same distribution as  $\{X_j : j \geq 1\}$  and the collections

$$\{X_1, \dots, X_n\} \text{ and } \{X_{n+m+1}, X_{n+m+2}, \dots\}$$

are independent.

**Lemma 2.1.** *Let  $\{X_n; n \geq 1\}$  be a sequence of stationary and  $m$ -dependent random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and  $\{a_{n,k}; k \geq 1, n \geq 1\}$  an array of non-negative real random variables on the same probability space independent of  $\{X_n; n \geq 1\}$  such that, for all  $\omega \in \Omega$ ,*

$$\sum_{k \geq 1} a_{n,k} = 1 \text{ for } n \geq 1.$$

If

$$\sup_{\omega \in \Omega} \sup_{k \geq 1} a_{n,k} \rightarrow 0 \text{ as } n \rightarrow \infty$$

and

$$\mathbb{E}(X_1 \vee 0) < \infty,$$

then for each  $n \geq 1$ ,  $\sum_{k \geq 1} a_{n,k} X_k$  is a well-defined  $[-\infty, \infty)$ -valued random variable with

$$\mathbb{E} \left( \sum_{k \geq 1} a_{n,k} X_k \right) = \mathbb{E} X_1$$

and

$$\sum_{k \geq 1} a_{n,k} X_k \rightarrow_p \mathbb{E} X_1.$$

*Proof.* We define the  $\sigma$ -algebra  $\mathcal{G}_n = \sigma(a_{n,k}; k \geq 1)$ . By conditioning on  $\mathcal{G}_n$  it is obvious that

$$\mathbb{E} \left( \sum_{k \geq 1} a_{n,k} X_k \right) = \mathbb{E} X_1.$$

For the second part, first assume that  $\mathbb{E}(X_1 \wedge 0) > -\infty$ . Then we know that

$$\mathbb{E}|X_1| < \infty$$

from the hypothesis that  $\mathbb{E}(X_1 \vee 0) < \infty$ . Given  $\epsilon > 0$  we choose  $\tau > 0$  such that

$$\mathbb{E}(|X_1|I_{\{|X_1| \geq \tau\}}) \leq \frac{\epsilon^2}{2}.$$

Now, for  $j \geq 1$ , define the following two random variables:

$$Y_j(\tau) = X_j I_{\{|X_1| \leq \tau\}} - \mathbb{E}(X_j I_{\{|X_1| \leq \tau\}}),$$

$$Z_j(\tau) = X_j - \mathbb{E}X_1 - Y_j(\tau).$$

Then because  $\sum_{k \geq 1} a_{n,k} = 1$  for all  $\omega \in \Omega$  and all  $n \geq 1$ , we have from the definition of  $Z_j(\tau)$ ,

$$\sum_{k \geq 1} a_{n,k} X_k - \mathbb{E}X_1 = \sum_{k \geq 1} a_{n,k} Y_k(\tau) + \sum_{k \geq 1} a_{n,k} Z_k(\tau).$$

Now

$$\begin{aligned} \mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} Z_k(\tau) \right| \geq \epsilon \middle| \mathcal{G}_n \right) &\leq \frac{1}{\epsilon} \mathbb{E} \left( \left| \sum_{k \geq 1} a_{n,k} Z_k(\tau) \right| \middle| \mathcal{G}_n \right) \\ &\leq \frac{1}{\epsilon} \sum_{k \geq 1} a_{n,k} \mathbb{E}(|Z_k(\tau)| \middle| \mathcal{G}_n). \end{aligned}$$

Now  $\mathbb{E}|Z_k(\tau)| < 2\mathbb{E}(|X_1|I_{\{|X_1| \geq \tau\}})$  by the definition of  $Z_k(\tau)$  so we use properties of conditional expectation to get

$$\begin{aligned} \mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} Z_k(\tau) \right| \geq \epsilon \right) &\leq \frac{2}{\epsilon} \mathbb{E}(|X_1|I_{\{|X_1| > \tau\}}) \\ &\leq \epsilon. \end{aligned}$$

We now use the stationarity and  $m$ -dependence of  $\{X_n; n \geq 1\}$  to get

$$\begin{aligned}
\mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} Y_k(\tau) \right| \geq \epsilon \middle| \mathcal{G}_n \right) &\leq \frac{1}{\epsilon^2} \text{Var} \left( \sum_{k \geq 1} a_{n,k} Y_k(\tau) \middle| \mathcal{G}_n \right) \\
&= \frac{1}{\epsilon^2} \left( \sum_{k \geq 1} a_{n,k}^2 \text{Var}(Y_k(\tau) | \mathcal{G}_n) \right. \\
&\quad \left. + 2 \sum_{1 \leq i-j \leq m} a_{n,i} a_{n,j} \text{Cov}(Y_i(\tau), Y_j(\tau) | \mathcal{G}_n) \right) \\
&\leq \frac{1}{\epsilon^2} (m+1) \sum_{k \geq 1} a_{n,k}^2 \text{Var}(Y_k(\tau) | \mathcal{G}_n) \\
&\leq \frac{2(m+1)\tau^2}{\epsilon^2} \sup_{k \geq 1} a_{n,k} \sum_{k \geq 1} a_{n,k} \\
&= \frac{2(m+1)\tau^2}{\epsilon^2} \sup_{k \geq 1} a_{n,k}.
\end{aligned}$$

Now remove the conditioning and use the condition on the  $a_{n,k}$  to obtain

$$\mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} Y_k(\tau) \right| \geq \epsilon \right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Hence

$$\limsup \mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} X_k - \mathbb{E}X_1 \right| \geq 2\epsilon \right) \leq \epsilon,$$

from which we can deduce the result, when  $\mathbb{E}(X_1 \wedge 0) > -\infty$ . If  $\mathbb{E}(X_1 \wedge 0) = -\infty$ , for each  $M > 0$ , we truncate and apply the previous argument to  $X_k I_{\{X_k \geq -M\}}$  as in Li and Rogers (1999).  $\square$

**Lemma 2.2.** *With the same framework as Lemma 2.1, and further that*

$$\zeta_n^{-1} := \sum_{k \geq 1} a_{n,k}^2 \leq cb^{-n} \text{ for all } n \geq 1, \omega \in \Omega$$

for some constants  $b > 1$  and  $c > 0$ , and

$$\mathbb{E}(|X_1| (L(|X_1|))^\delta) < \infty$$

for some  $\delta > 1$ , where  $L(x) := \log(\max\{e, x\})$ , then

$$\sum_{n \geq 1} \mathbb{P} \left( \left| \sum_{k \geq 1} a_{n,k} X_k - \mathbb{E}X_1 \right| \geq \epsilon \right) < \infty, \text{ for all } \epsilon > 0$$

and hence (via Borel-Cantelli) we have that

$$\lim_{n \rightarrow \infty} \sum_{k \geq 1} a_{n,k} X_k = \mathbb{E}X_1 \text{ a.s.}$$

*Proof.* By subtracting  $\mathbb{E}X_1$  from the initial random variables, we can assume  $\mathbb{E}X_1 = 0$ .

For each  $n \geq 1$ , define the following random variables:

$$U_n = \sum_{k \geq 1} a_{n,k} X_k I_{\{|X_k| \leq \zeta_n\}},$$

$$V_n = \sum_{k \geq 1} a_{n,k} X_k I_{\{|X_k| > \zeta_n\}},$$

and define  $\mathcal{G} = \sigma(a_{n,k}; n \geq 1, k \geq 1)$ . Now

$$\begin{aligned} \sum_{n \geq 1} \mathbb{P}(|V_n - \mathbb{E}(V_n | \mathcal{G})| \geq \epsilon | \mathcal{G}) &\leq \frac{1}{\epsilon} \sum_{n \geq 1} \mathbb{E}(|V_n - \mathbb{E}(V_n | \mathcal{G})| | \mathcal{G}) \\ &= \frac{1}{\epsilon} \sum_{n \geq 1} \mathbb{E}(|X_1 I_{\{|X_1| > \zeta_n\}} - \mathbb{E}(X_1 I_{\{|X_1| > \zeta_n\}} | \mathcal{G})| | \mathcal{G}) \\ &\quad \text{by the definition of } V_n. \\ &\leq \frac{2}{\epsilon} \sum_{n \geq 1} \mathbb{E}(|X_1 I_{\{|X_1| > \zeta_n\}}| | \mathcal{G}) \\ &\quad \text{by the triangle and Jensen's inequalities.} \end{aligned}$$

Now remove the conditioning to get

$$\begin{aligned} \sum_{n \geq 1} \mathbb{P}(|V_n - \mathbb{E}V_n| \geq \epsilon) &\leq \frac{2}{\epsilon} \sum_{n \geq 1} \mathbb{E}(|X_1 I_{\{|X_1| > b^n/c\}}|) \\ &\leq \frac{2}{\epsilon} \sum_{n=1}^{\infty} \sum_{j=n}^{\infty} \frac{b^{j+1}}{c} \mathbb{P}(b^j/c < |X_1| < b^{j+1}/c) \\ &\leq \frac{2}{\epsilon} \sum_{j=n}^{\infty} \frac{j b^{j+1}}{c} \mathbb{P}(b^j < |cX_1| < b^{j+1}) \\ &\leq \frac{2b}{\epsilon \log b} \mathbb{E}(|X_1| L(|cX_1|)) \\ &< \infty. \end{aligned}$$



Now we use variance and covariance arguments again to get

$$\begin{aligned} \mathbb{P}(|U_n - \mathbb{E}(U_n | \mathcal{G})| \geq \epsilon | \mathcal{G}) &\leq \frac{1}{\epsilon^2} \text{Var}(U_n | \mathcal{G}) \\ &\leq \frac{m+1}{\epsilon^2} \sum_{k \geq 1} a_{n,k}^2 \text{Var}(X_1 I_{\{|X_1| \leq \zeta_n\}} | \mathcal{G}) \\ &\leq \frac{m+1}{\epsilon^2} \zeta_n^{-1} \text{Var}(X_1 I_{\{|X_1| \leq \zeta_n\}} | \mathcal{G}). \end{aligned}$$

Now remove the conditioning to get

$$\begin{aligned} \sum_{n \geq 1} \mathbb{P}(|U_n - \mathbb{E}U_n| \geq \epsilon) &\leq \frac{m+1}{\epsilon^2} \sum_{n \geq 1} cb^{-n} \text{Var}(X_1 I_{\{|X_1| \leq b^n/c\}}) \\ &\leq \frac{m+1}{\epsilon^2} \sum_{n \geq 1} cb^{-n} \mathbb{E}(X_1^2 I_{\{|X_1| \leq b^n/c\}}) \\ &= \frac{m+1}{\epsilon^2} \sum_{n \geq 1} cb^{-n} \left( \frac{b^n}{c} O\left(\frac{1}{(L(\frac{b^n}{c}))^\delta}\right) \mathbb{E}(|X_1| (L(|X_1|))^\delta) \right) \\ &= \sum_{n \geq 1} O\left(\frac{1}{n^\delta}\right) \\ &< \infty. \end{aligned}$$

Combining the results for  $U_n$  and  $V_n$  we can now use the Borel-Cantelli Lemma to get the result.  $\square$

### 3. THE RANDOMIZED MODEL

We let  $\mathcal{X}^{(0)} = \{X_j^{(0)}; j \geq 1\}$  be a sequence of independent identically distributed random variables with  $\mathbb{P}(X^{(0)} \in \mathcal{D}) = 1$  for some closed domain  $\mathcal{D} \subseteq \mathbb{R}$ . We define  $\xi = \inf\{x : x \in \mathcal{D}\}$ . (Note that  $\xi$  may be  $-\infty$  but otherwise  $\xi \in \mathcal{D}$ .)

Now let  $Z_j^{(n)}; j \geq 1, n \geq 1$  be a set of independent (of each other and of  $\mathcal{X}^{(0)}$ ) and identically distributed random variables taking values in some measurable space  $R$ . For each  $r \in R$  let  $f_r : \mathcal{D}^k \rightarrow \mathcal{D}$  be a real measurable function of  $k$  variables, where  $k$  is a fixed integer,  $k > 1$ .

We now define a sequence  $\{\mathcal{X}^{(n)}; n \geq 0\}$  by

$$\begin{aligned} X_j^{(n)} &= f_{Z_j^{(n)}}(X_{(k-1)j+1}^{(n-1)}, X_{(k-1)j+2}^{(n-1)}, \dots, X_{kj}^{(n-1)}), \\ \mathcal{X}^{(n)} &= \mathcal{R}\mathcal{X}^{(n-1)} = \{X_j^{(n)}; j \geq 1\}. \end{aligned}$$

A model of this type, with  $R = \{1, 2\}$ ,  $f_1(x_1, x_2) = \frac{x_1+x_2}{2}$ ,  $f_2(x_1, x_2) = \min\{x_1, x_2\}$ , is discussed in Alon and Naor (1993); Boppana and Narayanan (1993, 2000) in connection with a ‘collective coin-flipping protocol’ and a ‘leader election protocol’.

Define  $\sigma$ -algebras  $\mathcal{F}_n = \sigma(X_i^{(j)}; 0 \leq j \leq n, i \geq 1)$  and  $\mathcal{G}_n = \sigma(Z_i^{(n)}; i \geq 1)$ .

The following result is based on Theorem 2.1(i) of Li and Rogers (1999) and the proof is largely the same except for the use of conditional expectation.

**Theorem 3.1.** *Suppose that we have non-negative constants  $\alpha_{i,r}$ ,  $1 \leq i \leq k, r \in R$ , such that*

$$\sum_{i=1}^k \alpha_{i,r} = 1 \text{ for all } r \text{ and } \alpha_{i,r} \leq A \text{ for all } i, r \text{ and some constant } A < 1.$$

*Further suppose that for all  $r \in R$ , we have the subadditivity constraint*

$$f_r(x_1, x_2, \dots, x_k) \leq \sum_{i=1}^k \alpha_{i,r} x_i.$$

*If*

$$\mathbb{E}(X_1^{(0)} \wedge 0) < \infty,$$

*then for some  $\lambda \in \mathcal{D} \cup \{\xi\}$ ,*

$$\mathbb{E}X_1^{(n)} \downarrow \lambda \text{ as } n \rightarrow \infty \text{ and } X_1^{(n)} \rightarrow_p \lambda.$$

*Proof.* Define

$$g_j^{(n)}(x_1, \dots, x_k) = \sum_{i=1}^k \alpha_{i, Z_j^{(n)}} x_i$$

and

$$\begin{aligned} h_j^{(n)}(x_1, \dots, x_{k^n}) &= g_j^{(n)}(h_{(j-1)k+1}^{(n-1)}(x_1, \dots, x_{k^{n-1}}), \\ &h_{(j-1)k+2}^{(n-1)}(x_{k^{n-1}+1}, \dots, x_{2k^{n-1}}), \dots, \\ &h_{jk}^{(n-1)}(x_{(k-1)k^{n-1}+1}, \dots, x_{k^n})) \end{aligned}$$

where  $h_j^{(0)}(x) = x$  for all  $j$ .

Note that

$$h_j^{(n)}(x_1, \dots, x_{k^n}) = \sum_{i=1}^{k^n} d_{i,j}^{(n)} x_i$$

where  $d_{i,j}^{(n)} = \alpha_{p,Z_j^{(n)}} d_{q,(j-1)k+q}^{(n-1)}$  where  $i = (p-1)k^{n-1} + q$  with  $1 \leq p \leq k$  and  $1 \leq q \leq k^{n-1}$ . Inductively, we see that  $\max_i d_{i,1}^{(n)} \leq A^n \rightarrow 0$  as  $n \rightarrow \infty$ . This allows us to use Lemma 2.1 (with  $a_{n,i} = d_{i,1}^{(n)}$ ) to give

$$\sum_{i=1}^{k^n} d_{i,1}^{(n)} X_i^{(0)} \rightarrow_p \mathbb{E}X_1^{(0)}. \quad (1)$$

Now

$$\begin{aligned} \mathbb{E}(X_1^{(n)} | \mathcal{F}_{n-1}, \mathcal{G}_n) &= f_{Z_1^{(n)}}(X_1^{(n-1)}, \dots, X_k^{(n-1)}) \\ &\leq \sum_{i=1}^k \alpha_{i,Z_1^{(n)}} X_i^{(n-1)} \end{aligned}$$

and using properties of conditional expectation and the i.i.d. property of  $\mathcal{X}^{(n-1)}$  we have

$$\begin{aligned} \mathbb{E}(X_1^{(n)} | \mathcal{G}_n) &\leq \sum_{i=1}^k \alpha_{i,Z_1^{(n)}} \mathbb{E}(X_i^{(n-1)} | \mathcal{G}_n) \\ &= \mathbb{E}(X_1^{(n-1)} | \mathcal{G}_n) \end{aligned}$$

but  $X_1^{(n-1)}$  is independent of  $\mathcal{G}_n$  so that  $\mathbb{E}X_1^{(n)}$ ,  $n \geq 1$  is a non-increasing sequence of real numbers, hence there exists  $\lambda \in [-\infty, \infty)$  such that

$$\mathbb{E}X_1^{(n)} \downarrow \lambda \text{ as } n \rightarrow \infty. \quad (2)$$

Then if  $\mathbb{E}X_1^{(0)} = -\infty$  we have, using (1) that  $X_1^{(n)} \rightarrow_p -\infty$  and if  $\mathbb{E}|X_1^{(0)}| < \infty$  then

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_1^{(n)} \geq \mathbb{E}X_1^{(0)} + \epsilon) = 0.$$

Further,

$$X_1^{(n)} I_{\{X_1^{(n)} \geq \mathbb{E}X_1^{(0)} + \epsilon\}} \leq \sum_{i=1}^{k^n} d_{n,i} |X_i^{(0)}| I_{\{\sum_{i=1}^{k^n} d_{n,i} X_i^{(0)} \geq \mathbb{E}X_1^{(0)} + \epsilon\}},$$

and for every  $M > 0$ ,

$$\begin{aligned} &\mathbb{E}\left(\sum_{i=1}^{k^n} d_{n,i} |X_i^{(0)}| I_{\{\sum_{i=1}^{k^n} d_{n,i} X_i^{(0)} \geq \mathbb{E}X_1^{(0)} + \epsilon\}}\right) \\ &\leq M \mathbb{P}\left(\sum_{i=1}^{k^n} d_{n,i} X_i^{(0)} \geq \mathbb{E}X_1^{(0)} + \epsilon\right) + \mathbb{E}\left(\sum_{i=1}^{k^n} d_{n,i} |X_i^{(0)}| I_{\{\sum_{i=1}^{k^n} d_{n,i} |X_i^{(0)}| \geq M\}}\right) \end{aligned}$$

which with  $\mathbb{E}|X_1^{(0)}| < \infty$  implies that

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_1^{(n)} I_{\{X_1^{(n)} \geq \mathbb{E}X_1^{(0)} + \epsilon\}}) = 0. \quad (3)$$

Now we repeat these arguments but starting with  $\mathcal{X}^{(n_0)}$  instead of with  $\mathcal{X}^{(0)}$ . The conditions on  $\mathcal{X}^{(0)}$  imply that  $\mathbb{E}(X_1^{(n_0)} \wedge 0) < \infty$  so they apply to  $\mathcal{X}^{(n_0)}$  as well. So if for any  $n_0$ ,  $\mathbb{E}X_1^{(n_0)} = -\infty$  then  $X_1^{(n)} \rightarrow_p -\infty$  and if  $\mathbb{E}|X_1^{(n_0)}| < \infty$  then

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_1^{(n)} \geq \mathbb{E}X_1^{(n_0)} + \epsilon) = 0$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_1^{(n)} I_{\{X_1^{(n)} \geq \mathbb{E}X_1^{(n_0)} + \epsilon\}}) = 0. \quad (4)$$

So if  $\lambda = -\infty$  then we have

$$X_1^{(n)} \rightarrow_p -\infty$$

and if  $\lambda > -\infty$ , for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_1^{(n)} \geq \lambda + \epsilon) = 0$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_1^{(n)} I_{\{X_1^{(n)} \geq \lambda + \epsilon\}}) = 0. \quad (5)$$

If  $\lambda > -\infty$  then

$$\begin{aligned} \mathbb{E}(X_1^{(n)} - \lambda) &= \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{X_1^{(n)} \geq \lambda + \epsilon'\}}\right) \\ &\quad + \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{X_1^{(n)} \leq \lambda - \epsilon\}}\right) + \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{\lambda - \epsilon < X_1^{(n)} < \lambda + \epsilon'\}}\right), \end{aligned}$$

so using (2) and (5) gives that

$$\begin{aligned} 0 &\geq \lim_{n \rightarrow \infty} \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{X_1^{(n)} \leq \lambda - \epsilon\}}\right) \\ &= - \lim_{n \rightarrow \infty} \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{\lambda - \epsilon < X_1^{(n)} < \lambda + \epsilon'\}}\right) \geq -\epsilon' \end{aligned}$$

for all  $\epsilon' > 0$ . Hence

$$\lim_{n \rightarrow \infty} \mathbb{E}\left((X_1^{(n)} - \lambda)I_{\{X_1^{(n)} \leq \lambda - \epsilon\}}\right) = 0 \quad (6)$$

and (5) and (6) give the result

$$X_1^{(n)} \rightarrow_p \lambda \text{ as } n \rightarrow \infty.$$

As in Li and Rogers (1999), the proof finishes by taking an a.s.-convergent subsequence to obtain  $\lambda \in \mathcal{D} \cup \{\xi\}$ .  $\square$

The next result is based on Theorem 2.1(ii) of Li and Rogers (1999) and gives a form of strong law, but only for the lim sup. Again the proof is based on that in Li and Rogers (1999).

**Lemma 3.2.** *With the hypotheses of Theorem 3.1 and additionally if  $\mathcal{D}$  is bounded below and we have*

$$\mathbb{E}(|X_1^{(0)}| (L(|X_1^{(0)}|))^{\delta}) < \infty$$

for some  $\delta > 1$  then for some  $\lambda \in \mathcal{D}$  the conclusion of Theorem 3.1 holds and additionally

$$\limsup_{n \rightarrow \infty} X_1^{(n)} = \lambda \text{ almost surely.}$$

*Proof.* With the hypotheses of Theorem 3.1 holding, we know that  $X_1^{(n)} \rightarrow_p \lambda$  for some  $\lambda \in \mathcal{D} \cup \{\xi\}$ . However  $\mathcal{D}$  is closed and bounded below, so  $\xi \in \mathcal{D}$ , hence  $\lambda \in \mathcal{D}$ . There exists a subsequence  $\{n_j; j \geq 1\}$  such that

$$\lim_{j \rightarrow \infty} X_1^{(n_j)} = \lambda \text{ a.s.}$$

and hence

$$\limsup_{j \rightarrow \infty} X_1^{(n_j)} \geq \lambda \text{ a.s.} \quad (7)$$

Now  $X_1^{(n)} \leq \sum_{i=1}^{k^n} d_{i,1}^{(n)} X_i^{(0)}$  as before. The condition  $\sum_{i=1}^{k^n} (d_{i,1}^{(n)})^2 \leq cb^{-n}$  ( $b > 1, c > 0$ ) necessary to apply Lemma 2.2 is easy in our case (see Example 2.1 of Li and Rogers (1999)). So we have

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{k^n} d_{i,1}^{(n)} X_i^{(0)} = \mathbb{E}X_1^{(0)} \text{ a.s.}$$

and hence

$$\limsup_{n \rightarrow \infty} X_1^{(n)} \leq \mathbb{E}X_1^{(0)} \text{ a.s.} \quad (8)$$

As in the proof of Theorem 3.1, we now observe that the conditions on  $\mathcal{X}^{(0)}$  imply the same conditions on  $\mathcal{X}^{(n_0)}$  for any integer  $n_0$  (we need  $\mathcal{D}$  to be bounded below here, and also Jensen's inequality). So we can use identical arguments to those above to obtain

$$\limsup_{n \rightarrow \infty} X_1^{(n)} \leq \mathbb{E}X_1^{(n_0)} \text{ a.s.} \quad (9)$$

and letting  $n_0 \rightarrow \infty$ ,

$$\limsup_{n \rightarrow \infty} X_1^{(n)} \leq \lambda \text{ a.s.} \quad (10)$$

Then (7) and (10) give the result.  $\square$

The next result is based on Theorem 2.1(iii) of Li and Rogers (1999) although the symmetry condition is slightly modified. It gives strong-law convergence in a restricted case.

**Lemma 3.3.** *Assume the hypotheses of Lemma 3.2, and further the following symmetry condition:*

*For all  $\sigma \in S_k$ , the permutation group on  $k$  elements, we have*

$$\mathbb{P}(f_{Z_i^{(n)}} \in E) = \mathbb{P}(f_{Z_i^{(n)}} \circ \sigma \in E)$$

*for all  $\mathbb{P}$ -measurable subsets  $E$  of measurable functions of  $k$  real variables (i.e. changing the order of the input variables does not alter the conditional distributions of  $\mathcal{X}^{(n-1)}$  and  $\mathcal{X}^{(n)}$ ).*

*Then for some  $\lambda \in \mathcal{D}$ , we have*

$$\lim_{n \rightarrow \infty} X_1^{(n)} = \lambda \text{ a.s.}$$

*Further, if  $\alpha_{i,r} = 1/k$  for all  $i, r$ , then the conclusion holds under the weaker condition that  $\mathbb{E}|X_1^{(0)}| < \infty$ .*

*Proof.* The hypotheses of Lemma 3.2 apply so we know that there exists a  $\lambda \in \mathcal{D}$  such that

$$X_1^{(n)} \rightarrow_p \lambda$$

and

$$\limsup X_1^{(n)} = \lambda \text{ a.s.} \quad (11)$$

Similarly for every  $l > 1$  we have  $X_l^{(n)} \rightarrow_p \lambda$  and  $\limsup X_l^{(n)} = \lambda$  a.s.

Now set

$$Y^{(n)} = \sum_{i=1}^k X_i^{(n)}.$$

For  $n \geq 0$ , define  $\sigma$ -algebras

$$\mathcal{F}'_n = \sigma(Y^{(l)}; l \geq n).$$

Our symmetry condition gives that, for  $1 \leq i \leq k$ ,

$$\mathbb{E}(X_{(i-1)k+1}^{(n)} + \cdots + X_{ik}^{(n)} | \mathcal{F}'_{n+1}) = \mathbb{E}(Y^{(n)} | \mathcal{F}'_{n+1})$$

and hence

$$\mathbb{E}(Y^{(n)} | \mathcal{F}'_{n+1}) = \frac{\mathbb{E}(X_1^{(n)} + \cdots + X_{k^2}^{(n)} | \mathcal{F}'_{n+1})}{k}.$$

We note that the symmetry condition implies that  $\mathbb{E}\alpha_{i,Z_j^{(n+1)}} = 1/k$  and further that  $Z_j^{(n+1)}$  is independent of  $\mathcal{F}'_{n+1}$ , and so

$$\mathbb{E}(\alpha_{i,Z_j^{(n+1)}} | \mathcal{F}'_{n+1}) = \frac{1}{k}.$$

Hence conditioning on the  $Z_j^{(n+1)}$  and using the subadditivity of the functions will give

$$\mathbb{E}\left(\frac{X_1^{(n)} + \cdots + X_k^{(n)}}{k} \middle| \mathcal{F}'_{n+1}\right) \geq \mathbb{E}(X^{(n+1)} | \mathcal{F}'_{n+1})$$

and similarly for the other  $X_k^{(n+1)}$ , giving

$$\begin{aligned} \mathbb{E}(Y^{(n)} | \mathcal{F}'_{n+1}) &\geq \mathbb{E}(X_1^{(n+1)} + \cdots + X_k^{(n+1)} | \mathcal{F}'_{n+1}) \\ &= Y^{(n+1)} \text{ a.s.} \end{aligned}$$

Hence the sequence  $\{Y^{(n)}; n \geq 1\}$  is a reversed time submartingale, so as in Li and Rogers (1999) we can use the convergence theorem and the convergence in probability to conclude that

$$\liminf_{n \rightarrow \infty} X_1^{(n)} \geq \lambda \text{ a.s.} \quad (12)$$

Then (11) and (12) give the result.

When  $\alpha_{i,r} = 1/k$  we can use the Strong Law of Large Numbers instead of Lemma 2.2 (as in the proof of Theorem 2.1(iii) of Li and Rogers (1999)) and so the final part of the statement follows.  $\square$

The following result and its proof are based on Corollary 3.1 of Li and Rogers (1999).

**Corollary 3.4.** *For any set of random variables  $\tilde{X}^{(n)}$  where  $\tilde{X}^{(n)}$  has the same distribution as  $X_1^{(n)}$  for all  $n$ ,*

$$\tilde{X}^{(n)} \rightarrow \lambda \text{ a.s.}$$

*Proof.* For each  $n \geq 1$ ,  $X_2^{(n)}$  is determined by the random variables  $\{X_i^{(0)}; k^n + 1 \leq i \leq 2k^n\} \cup \{Z_i^{(j)}; 1 \leq j \leq n; k^{n-j} + 1 \leq i \leq 2k^{n-j}\}$ . But these sets are independent for different  $n$  and so  $\{X_2^{(n)}; n \geq 0\}$  is a set of independent random variables. From the proof of Lemma 3.3

we have that  $X_2^{(n)} \rightarrow \lambda$  a.s. and so by Borel-Cantelli we have that

$$\sum_{n=0}^{\infty} \mathbb{P}(|X_2^{(n)} - \lambda| \geq \epsilon) < \infty \text{ for all } \epsilon > 0. \quad (13)$$

But, for each  $n$ ,  $X_2^{(n)}$  and  $X_1^{(n)}$  are i.i.d. random variables. By the hypotheses of the corollary,  $\tilde{X}^{(n)}$  has the same distribution as  $X_2^{(n)}$  and hence (13) implies that

$$\sum_{n=0}^{\infty} \mathbb{P}(|\tilde{X}^{(n)} - \lambda| \geq \epsilon) < \infty \text{ for all } \epsilon > 0,$$

which, using Borel-Cantelli, gives the result.  $\square$

We can now use this to strengthen Lemma 3.3 by removing the symmetry condition. Theorem 1.2 is a special case of the following:

**Theorem 3.5.** (a) *Under the hypotheses of Lemma 3.2 there exists  $\lambda \in \mathcal{D}$  with*

$$X_1^{(n)} \rightarrow \lambda \text{ a.s.}$$

(b) *If  $\alpha_{i,r} = 1/k$  for all  $i, r$  then the conclusion holds under the weaker condition that  $\mathbb{E}|X_1^{(0)}| < \infty$ .*

*Proof.* We define

$$g_r(x_1, \dots, x_k) = \sum_{i=1}^k \alpha_{i,r} x_i \text{ for all } (x_1, \dots, x_k) \in \mathcal{D}^k.$$

Set  $\tilde{R} = R \times S_k$  where  $S_k$  is the  $k$ th permutation group. Then, for each  $r, \sigma$ ;  $r \in R, \sigma \in S_k$ , define  $\tilde{f}_{(r,\sigma)} : \mathcal{D}^k \rightarrow \mathcal{D}$  by

$$\tilde{f}_{(r,\sigma)}(x_1, \dots, x_k) = f_r \circ \sigma(x_1, \dots, x_k).$$

The condition on the  $f_r$  ensures that, for all  $r$ ,

$$\begin{aligned} \tilde{f}_{(r,\sigma)}(x_1, \dots, x_k) &\leq \tilde{g}_{(r,\sigma)}(x_1, \dots, x_k) \\ &:= g_r \circ \sigma(x_1, \dots, x_k) \text{ for all } (x_1, \dots, x_k) \in \mathcal{D}^k. \end{aligned}$$

We now consider a model with, with random variables  $Y_j^{(n)}$ , based on the set of functions  $\{\tilde{f}_{\tilde{r}}; \tilde{r} \in \tilde{R}\}$ . The symmetry condition of Lemma 3.3 will be satisfied if we define the random variables  $\tilde{Z}_i^{(n)}$  so that  $\mathbb{P}(\tilde{Z}_i^{(n)} \in E \times \{\sigma\}) = (1/k!) \mathbb{P}(Z_i^{(n)} \in E)$  for measurable subsets  $E \subseteq R$



and  $\sigma \in S_k$ . So if we set  $\mathcal{Y}^{(0)} = \mathcal{X}^{(0)}$  and define i.i.d. sequences  $\mathcal{Y}^{(n)} = \{Y_i^{(n)} : i \geq 1\}$  using our new model, we can apply Lemma 3.3 to obtain  $Y_1^{(n)} \rightarrow \lambda$  a.s. for some  $\lambda \in \mathcal{D}$ .

Now note that, as  $\{Y_i^{(n)} : i \geq 1\}$  are i.i.d., the distribution of  $Y_i^{(1)}$  is independent of  $Z_i^{(1)}$ . Hence it is the same as the distribution of  $X_i^{(1)}$ . We continue inductively to see that  $X_i^{(n)}$  and  $Y_i^{(n)}$  have the same distribution for all  $n, i$  (although the joint distributions as  $n$  varies will not necessarily be the same). We now apply Corollary 3.4 to obtain the result.

To apply Lemma 3.3 we need  $\mathbb{E} \left( |X_1^{(0)}| (L(|X_1^{(0)}|))^{\delta} \right) < \infty$ , giving (a), except when  $\alpha_i = 1/k$  when  $\mathbb{E}|X_1^{(0)}| < \infty$  is sufficient, giving (b).  $\square$

Theorem 1.2 is just Theorem 3.5 in the case  $R = \{1\}$ .

It is conjectured in Li and Rogers (1999) that an almost sure limit will exist for more general iterations which may vary from stage to stage, and with the initial random variables  $m$ -dependent rather than independent. In the case where  $k$  varies the reversed submartingale argument of this paper cannot be used and the argument in Theorem 3.5 requires independence initially, so the methods used here cannot be applied in their present form to this conjecture.

## REFERENCES

- Alon, N. and Naor, M. (1993), Coin-flipping games immune against linear sized coalitions, *SIAM J. Comput.* **22** 403-417.
- Alon, N. and Rabin, M. O. (1989), Biased coins and randomized algorithms, in *Advances in Computing Research 5: Randomness and Computation* 499-507, S. Micali, ed., JAI Press, Greenwich, Connecticut.
- Bernasconi, J. (1978), Real renormalization of bond disordered conductance lattices, *Phys. Rev. B* **18** 2185-2191.
- Blumenfeld, R. (1988), Probability densities of homogeneous functions: Explicit approximations and applications to percolating networks, *J. Phys. A* **21** 815-825.

- Boppana, R. and Narayanan, B. (1993), The biased coin problem, in *Proceedings of the 35th ACM Symposium on Theory of Computing* 252-257, ACM Press, New York.
- Boppana, R. and Narayanan, B. (2000), Perfect-information leader election with optimal resilience, *SIAM J. Comput.* **29** 1304-1320.
- Derrida, R. (1986), Pure and random models of statistical mechanics on hierarchical lattices, in *Critical Phenomena, Random Systems and Gauge Theories* 989-999, ed. K. Osterwalder and R. Stora, North-Holland, Amsterdam.
- Essoh, C. D. and Bellisard, J. (1989), Resistance and fluctuation of a fractal network of random resistors: a non-linear law of large numbers, *J. Phys. A* **22** 4537-4548.
- Griffiths, R. and Kaufman, M. (1982), Spin systems on hierarchical lattices: Introduction and thermodynamic limit, *Phys. Rev. B* **26** 5022-5032.
- Hambly, B. and O'Connell, N. (2000), A law of large numbers for random hierarchical sequences, BRIMS Technical Report HPL-BRIMS-2000-13.
- Jona-Lasinio, J. (1975), The renormalization group: A probabilistic view, *Nuovo Cimento B* **26** 99-119.
- Li, D. L. and Rogers, T. D. (1999), Asymptotic behavior for iterated functions of random variables, *Ann. App. Prob.* **9** 1175-1201.
- Newman, W. I., Gabrielov, A. M., Durand, T. A., Phoenix, S. L. and Turcotte, D. L. (1994), An exact renormalization model for earthquakes and material failure, *Physica D* **77** 200-216.
- Schenkel, A., Wehr, J. and Wittwer, P. (2000), Computer-assisted proofs for fixed point problems in Sobolev spaces, *Elec. J. Math. Phys.* **6** No. 3.
- Schlösser, T. and Spohn, H. (1992), Sample-to-sample fluctuations in the conductivity of a disordered medium, *J. Stat. Phys.* **69** 955-967.
- Shneiberg, I. Ya. (1986), Hierarchical sequences of random variables, *Theory Probab. Appl.* **31** 137-141.
- Stinchcombe, R. B. and Watson, B. P. (1976), Renormalization group approach for percolation conductivity, *J. Phys. C* **9** 3221-3247.
- Wehr, J. (1997), A strong law of large numbers for iterated functions of independent random variables, *J. Stat. Phys.* **86** 1373-1384.
- Wehr, J. and Woo, J-M. (2000), Random conductance on hierarchical lattice, preprint.

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, UNIVERSITY WALK, BRISTOL BS8 1TW, AND, BRIMS, HEWLETT-PACKARD LABS, FILTON ROAD, STOKE GIFFORD, BS34 8QZ

*E-mail address:* `jonathan.jordan@bristol.ac.uk`