Journal of Statistical Physics, Vol. 104, Nos. 3/4, 2001

Central Limit Theorems for Nonlinear Hierarchical Sequences of Random Variables

Jan Wehr[1] **and Jung M. Woo**[1]

November 13, 2000; revised March 12, 2001

Let a random variable x_0 and a function $f: [a, b]^k \to [a, b]$ be given. A hierarchical sequence $\{x_n: n = 0, 1, 2,...\}$ of random variables is defined inductively by the relation $x_n = f(x_{n-1,1}, x_{n-1,2}..., x_{n-1,k})$, where $\{x_{n-1,i}: i = 1, 2,..., k\}$ is a family of independent random variables with the same distribution as x_{n-1}. We prove a central limit theorem for this hierarchical sequence of random variables when a function f satisfies a certain averaging condition. As a corollary under a natural assumption we prove a central limit theorem for a suitably normalized sequence of conductivities of a random resistor network on a hierarchical lattice.

KEY WORDS: Random resistor networks; central limit theorem; hierarchical lattices; renormalization group.

1. INTRODUCTION

The classical central limit theorem for independent and identically distributed (shortly IID) random variables can be presented in the following way: let x_0 be a random variable with a finite variance. We define inductively a sequence $\{x_n: n = 0, 1, 2,...\}$ of random variables as follows. Given x_{n-1}, let $x_{n-1,1}$ and $x_{n-1,2}$ be independent random variables with the same distribution as x_{n-1} and define

$$x_n = \frac{x_{n-1,1} + x_{n-1,2}}{2}. \tag{1}$$

[1] Department of Mathematics and Program in Applied Mathematics, University of Arizona, Tucson, Arizona 85721; e-mail: wehr@math.arizona.edu and jwoo@math.arizona.edu

0022-4715/01/0800-0777$19.50/0

The central limit theorem implies that the normalized random variables

$$\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}}$$

converge in distribution to a unit normal variable. The question we address is when does an analogous statement hold for sequences defined in a way similar to (1) where the arithmetic mean is replaced by a more general function f of two or more independent copies of x_{n-1},

$$x_n = f(x_{n-1,1}, x_{n-1,2}, \ldots, x_{n-1,k}). \tag{2}$$

An example of a hierarchical sequence of this type is the sequence of conductivities arising in a random resistor network on the so-called diamond hierarchical lattice. In this case $k = 4$ and f in the expression (2) is given by

$$f(u_1, u_2, u_3, u_4) = \frac{1}{\dfrac{1}{u_1} + \dfrac{1}{u_2}} + \frac{1}{\dfrac{1}{u_3} + \dfrac{1}{u_4}}. \tag{3}$$

The above function f expresses the effective conductivity of two sequential connections of conductors which are connected in parallel, where u_i $(i = 1, 2, 3, 4)$ represent the conductivities of the individual conductors. It is well defined for $u_1, \ldots, u_4 \geqslant 0$, including the case when one or more of the conductors is a perfect insulator (i.e., when one or more of the u_i equals zero). This model has been studied extensively in the physics literature as a simple model for conductivity of a two-dimensional random medium (see ref. 6 and references therein). Iterating the function f can be thought of as a simple model of the renormalization group map.

In this paper we prove the central limit theorem for a hierarchical sequence $\{x_n\}$ of random variables given by (2) when the function f satisfies a certain averaging condition (see Definition 1 and Theorem 2 below).

Definition 1. Let a and b two real numbers with $a < b$. We will say that a continuous function $f\colon [a, b]^k \to [a, b]$ is averaging if the following three conditions hold:

1. For all $u_i \in [a, b]$ $(i = 1, 2, \ldots, k)$

$$\min_i u_i \leqslant f(u_1, u_2, \ldots, u_k) \leqslant \max_i u_i.$$

2. f is monotone increasing, that is, for all u_i and u_i' (with $u_i \leqslant u_i'$ for $i = 1, 2,..., k$)

$$f(u_1, u_2,..., u_k) \leqslant f(u_1', u_2',..., u_k').$$

3. For all $u < v$ in $[a, b]$ and for any two distinct indices $i_1, i_2 \in \{1, 2,..., k\}$ there exist $u_i \in \{u, v\}$ $(i = 1, 2,..., k)$ such that $u_{i_1} = u$, $u_{i_2} = v$ and

$$u < f(u_1, u_2,..., u_k) < v.$$

Remark. Assumptions 1 and 2 are natural requirements on any averaging operation. Assumption 3 says that f is sensitive enough to varying its variables. This property is satisfied by several f which occur naturally in applications. An example of an f which satisfies conditions 1 and 2 but not condition 3 is $f(u_1,..., u_k) = \min(u_1,..., u_k)$. As is well known, the hierarchical sequence defined by iterating this f satisfies a non-Gaussian limit theorem (see ref. 2, p. 85).

Theorem 2. Let $a < b$. Let $f\colon [a, b]^k \to [a, b]$ be averaging. Assume that the range of a random variable x_0

$$\mathscr{R}(x_0) = \{u\colon \text{ for any } \epsilon > 0 \ \mathbf{P}[|x_0 - u| < \epsilon] > 0\}$$

is contained in $[a, b]$ and consists of more than one point. Suppose there exists $c \in (a, b)$ such that a hierarchical sequence x_n, defined as in (2), converges to c in probability. Also assume that f is twice continuously differentiable in the neighborhood of $(c, c,..., c)$ and that $\frac{\partial f}{\partial u_i}(c, c,..., c) \neq 0$ (and hence > 0 by Assumption 2 above) for at least two distinct indices i. Then the random variables

$$\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}}$$

converge to a unit normal variable in distribution.

A central limit theorem for the conductivity of a class of random resistor networks on hierarchical lattices will be proven as a corollary of Theorem 2 (Corollary 4).

The paper is organized as follows. In Section 2 we introduce random resistor networks on hierarchical lattices as an example of iterations of averaging functions. An application of Theorem 2 to this particular situation gives a central limit theorem for conductivity of a class of random resistor networks (Corollary 4). In Section 3 we state and prove a general central limit theorem (Proposition 5) for sequences in which the $n+1$-st term has the same distribution as a linear combination of k independent copies of the nth term up to a small (in a sense we make precise) correction. This theorem is used in Section 4 to prove our main result (Theorem 2). Proofs of two technical results used in Section 2 (Proposition 3) and in Section 4 (Lemma 8) are put in the appendix to increase readability of the main text of the paper.

2. RANDOM RESISTOR NETWORKS

We start from a brief discussion of random resistor networks on hierarchical lattices. More details can be found in refs. 3 and 6. Let $\mathbf{G} = (\mathscr{S}, \mathscr{B}, (s_t, s_b))$ be a connected graph with the set of sites (vertices) $\mathscr{S}$, the set of bonds (edges) $\mathscr{B}$, the top site s_t and the bottom site s_b. The top and bottom sites are called surface sites. These are the sites we apply a unit potential difference to. All other sites are called internal sites. Let k be the total number of bonds of $\mathbf{G}$. We will assume that any self-avoiding path connecting the two surface sites of $\mathbf{G}$ has at least length 2, and that there are at least two bond-disjoint self-avoiding paths connecting the two surface sites of $\mathbf{G}$.

Let us label the bonds of $\mathbf{G}$ by integers $1,\ldots, k$. For each $i = 1,\ldots, k$ let $u_i \geqslant 0$ be the conductivity of the ith bond of $\mathbf{G}$ and let $h(u_1, u_2,\ldots, u_k)$ denote the effective conductance of the resulting resistor network between its surface sites. In other (perhaps more familiar) words, we can introduce the resistivities $\frac{1}{u_i}$ and calculate the effective resistance of the system of resistors using Kirchhoff laws. h is then the inverse of the effective resistance. Define $f\colon [0, \infty)^k \to [0, \infty)$ by

$$f(u_1, u_2,\ldots, u_k) = \frac{h(u_1, u_2,\ldots, u_k)}{h(1, 1,\ldots, 1)}. \tag{4}$$

f is the effective conductivity (normalized conductance) of the resistor network.

Proposition 3. f given by (4) is an averaging function in the sense of Definition 1. Proof is given in the Appendix.

Given a nonnegative random variable x_0, representing the conductivity of an individual bond of our graph, we now obtain, by iterating f as in (2), a hierarchical sequence x_n of random variables, which can be interpreted as effective conductivities of a sequence of graphs $\mathbf{H_n}$ defined as follows.

A hierarchical lattice $\mathbf{H} = \{\mathbf{H_n}\colon n = 0, 1, 2,\ldots\}$ is generated inductively from a fixed graph $\mathbf{G}$ with k bonds, satisfying the above assumptions. At the level 0 $\mathbf{H_0}$ is a single bond, at the level 1 $\mathbf{H_1} = \mathbf{G}$, and at any level $n > 1$ a graph $\mathbf{H_n}$ is constructed by replacing each bond of $\mathbf{G}$ by a copy of $\mathbf{H_{n-1}}$, whose surface sites are placed at the endpoints of the bond. Hence $\mathbf{H_n}$ is a graph with k^n bonds, to the two surface sites of which a potential difference is applied. The (random) effective conductivity of $\mathbf{H_n}$ (cf. (4)) between the surface sites is denoted by x_n.

We are interested in the fluctuations of the effective conductivity of the hierarchical lattice in the infinite volume limit ($n \to \infty$). Corollary 4 is a central limit theorem for the conductivities x_n under the assumption that the limiting effective conductivity does not vanish, i.e., that the system is in the conducting phase. More precisely, let μ_0 be a distribution of x_0. In ref. 7, Shneiberg proved that there exists a $c(\mu_0)$ such that x_n converges to $c(\mu_0)$ in probability (see ref. 4 for related results[2]). Moreover, there exists a number $p_c(\mathbf{H})$ such that

$$c(\mu_0) = \begin{cases} 0 & \text{if} \quad \mathbf{P}[x_0 > 0] \leqslant p_c(\mathbf{H}) \\ \text{positive} & \text{if} \quad \mathbf{P}[x_0 > 0] > p_c(\mathbf{H}). \end{cases}$$

We mention here (though we will not use this) that $p_c(\mathbf{H})$ is the critical density of bond percolation on $\mathbf{H}$. Our assumptions on $\mathbf{G}$, that is, that any self-avoiding path connecting the two surface sites has length at least two and that there exist at least two bond-disjoint self-avoiding paths connecting the surface sites, guarantee that $0 < p_c(\mathbf{H}) < 1$. The above formula says that the system is in a conducting phase ($c(\mu_0) > 0$) if and only if it is in a supercritical phase ($\mathbf{P}[x_0 > 0] > p_c(\mathbf{H})$). Note also that if the system is in a conducting phase then f satisfies all the assumptions of Theorem 2 and we obtain the following corollary:

Corollary 4. Let b be a positive number. Let x_n be the effective conductivity of a random resistor network on a hierarchical lattice $\mathbf{H} = \{\mathbf{H_n}: n = 0, 1, 2,...\}$ described above. Assume that the range of x_0 is in the interval $[0, b]$ and consists of more than one point. Let $p_c(\mathbf{H})$ be a critical density of bond percolation on $\mathbf{H}$. If $\mathbf{P}[x_0 > 0] > p_c(\mathbf{H})$ then the random variables

$$\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}}$$

converge in distribution to a unit normal variable.

Remark on Corollary 4. Let $p = \mathbf{P}[x_0 > 0]$ and $g(p) = \mathbf{P}[x_1 > 0]$. Note that the function $g(p)$ is a probability that the two surface sites of $\mathbf{G}$

[2] See also errata at the end of this volume.

are connected by a set of bonds consisting of positive conductors, hence does not depend on the distribution μ_0 of x_0 but depends only on the parameter p. The exact value of $p_c(\mathbf{H})$ can be found as a unique repulsive fixed point of the equation $g(p) = p$ in the interval $(0, 1)$ (see ref. 7). In case of diamond lattice f is given by (3), $g(p) = 2p^2 - p^4$, hence $p_c(\mathbf{H}) = \frac{\sqrt{5}-1}{2}$.

Corollary 4 says that if the system is in a conducting phase then a sequence of suitably normalized effective conductivities converges to a unit normal variable in distribution. At the critical point $(\mathbf{P}[x_0 > 0] = p_c(\mathbf{H}))$ we expect a non-Gaussian behavior for x_n. In case of the diamond hierarchical lattice, this non-Gaussian behavior is studied in ref. 5.

3. A CENTRAL LIMIT THEOREM FOR A CLASS OF SEQUENCES SATISFYING AN APPROXIMATE LINEAR RECURSION

Our main tool for proving Theorem 2 is the following general central limit theorem which may be of independent interest (e.g., in study of conductivity problems on translationally invariant lattices).

Proposition 5. For all $n \geqslant 0$, let k be a positive integer greater than or equal to 2. Assume that a sequence $\{x_n : n = 0, 1, 2,...\}$ of real-valued random variables satisfies the following recursive relation.

$$x_{n+1} = \sum_{i=1}^{k} \alpha_{n,i} x_{n,i} + z_n,$$

where for each $n \geqslant 0$ $\alpha_{n,i}$ are real numbers, z_n is a real-valued random variable, and $\{x_{n,i} : i = 1,..., k\}$ are IID random variables with same distribution as x_n. Let us assume that for each $i \leqslant k$ $\alpha_{n,i}$ converges to some real number α_i. We also assume that $\alpha_i \neq 0$ for at least two distinct indices i. Let $\lambda_n = \sqrt{\sum_{i=1}^{k} \alpha_{n,i}^2}$ and $\lambda = \lim_{n \to \infty} \lambda_n$. Furthermore, assume that there exist $\delta_1 > 0$, $\delta_2 > 0$, $C_1 > 0$ and $C_2 > 0$ with $\delta_1 < \delta_2$ such that for any n

$$\mathbf{Var}[x_n] \geqslant C_1^2 \lambda^{2n} (1 - \delta_1)^{2n}$$
$$\mathbf{Var}[z_n] \leqslant C_2^2 \lambda^{2n} (1 - \delta_2)^{2n}.$$

Then the random variables

$$\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}}$$

converge in distribution to a unit normal variable.

Let us start by introducing some notation. For all $n \geqslant 0$ let

$$y_n = \sum_{i=1}^{k} \alpha_{n,i} x_{n,i} \tag{5}$$

and for all $n \geqslant 1$ let

$$\begin{aligned} \tilde{x}_n &= \frac{x_n - \mathbf{E}[x_n]}{\lambda_0 \lambda_1 \cdots \lambda_{n-1}} \\ \tilde{y}_n &= \frac{y_n - \mathbf{E}[y_n]}{\lambda_0 \lambda_1 \cdots \lambda_{n-1}} \\ \tilde{z}_n &= \frac{z_n - \mathbf{E}[z_n]}{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}. \end{aligned} \tag{6}$$

We then have

$$\tilde{y}_n = \sum_{i=1}^{k} \alpha_{n,i} \tilde{x}_{n,i}. \tag{7}$$

Then the assumptions in Proposition 5 say that for all $n \geqslant 1$

$$\begin{aligned} \tilde{x}_{n+1} &= \frac{1}{\lambda_n} (\tilde{y}_n + \tilde{z}_n) \\ \sqrt{\mathbf{Var}[\tilde{y}_n]} &= \lambda_n \sqrt{\mathbf{Var}[\tilde{x}_n]} \\ \sqrt{\mathbf{Var}[\tilde{x}_n]} &\geqslant C_1 (1-\delta_1)^n \\ \sqrt{\mathbf{Var}[\tilde{z}_n]} &\leqslant C_2 (1-\delta_2)^n. \end{aligned} \tag{8}$$

Proposition 5 will be proved using the fact that the characteristic functions of $\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}}$ converge to that of a Gaussian. The following lemma, which studies convergence of variances of $\tilde{x}_n$ is a step in this direction. Note that

$$\frac{x_n - \mathbf{E}[x_n]}{\sqrt{\mathbf{Var}[x_n]}} = \frac{\tilde{x}_n}{\sqrt{\mathbf{Var}[\tilde{x}_n]}}. \tag{9}$$

Lemma 6. The limit

$$\sigma_\infty = \lim_{n \to \infty} \sqrt{\mathbf{Var}[\tilde{x}_n]}$$

exists and is a positive real number. Moreover, there exist positive real constants C_3, C_4, and C_5 such that for all $n \geqslant 1$

$$\sqrt{\mathbf{Var}[\tilde{x}_n]} \geqslant C_3$$

$$\sqrt{\mathbf{Var}[\tilde{x}_n]} \leqslant C_4$$

$$\sqrt{\mathbf{Var}[\tilde{y}_n]} \leqslant C_5$$

Proof. From (8), using the triangle inequality, we get

$$\begin{aligned} |\sqrt{\mathbf{Var}[\tilde{x}_{n+1}]} - \sqrt{\mathbf{Var}[\tilde{x}_n]}| &= \left| \sqrt{\mathbf{Var}[\tilde{x}_{n+1}]} - \frac{1}{\lambda_n} \sqrt{\mathbf{Var}[\tilde{y}_n]} \right| \\ &\leqslant \frac{1}{\lambda_n} \sqrt{\mathbf{Var}[\tilde{z}_n]} \\ &\leqslant \frac{1}{\underline{\lambda}} C_2 (1-\delta_2)^n \end{aligned}$$

where $\underline{\lambda} = \inf_n \lambda_n > 0$. By summing up the above inequality over all $n \geqslant m$, we obtain a positive constant C_6 such that for all $m \geqslant 1$

$$\sum_{n=m}^{\infty} |\sqrt{\mathbf{Var}[\tilde{x}_{n+1}]} - \sqrt{\mathbf{Var}[\tilde{x}_n]}| \leqslant C_6 (1-\delta_2)^m.$$

Since the above series converges, existence of (finite) σ_∞ follows. Moreover, for all $m \geqslant 1$

$$|\sigma_\infty - \sqrt{\mathbf{Var}[\tilde{x}_m]}| \leqslant C_6 (1-\delta_2)^m,$$

which, in view of (8) implies that

$$C_1 (1-\delta_1)^m \leqslant \sqrt{\mathbf{Var}[\tilde{x}_m]} \leqslant \sigma_\infty + C_6 (1-\delta_2)^m.$$

It now follows from the assumption that $\delta_2 > \delta_1$, that σ_∞ is not zero. The rest of Lemma 6 follows immediately from the positivity of σ_∞.

The next lemma will allow us to extend the result about convergence of the variances to that of the characteristic functions. Its content are elementary characteristic function estimates.

Lemma 7. Let C_7 be any positive constant, and let X and Y be random variables with zero means and variances less than C_7^2. Then for all $t(|t| < \frac{1}{C_7})$

$$|\ln \mathbf{E}[\exp(itX)] - \ln \mathbf{E}[\exp(itY)]| \leqslant 4C_7 t^2 \sqrt{\mathbf{E}[(X-Y)^2]}.$$

Also we have

$$\lim_{t \to 0} \frac{\ln \mathbf{E}[\exp(itX)]}{t^2} = -\frac{1}{2} \mathbf{Var}[X].$$

Proof. Since $|e^{iu} - 1 - iu| \leqslant \frac{u^2}{2}$ for any real number u,

$$|\mathbf{E}[\exp(itX)] - 1| \leqslant t^2 \frac{\mathbf{E}[X^2]}{2}.$$

Therefore for $|t| < \frac{1}{C_7}$, the value $\mathbf{E}[\exp(itX)]$ lies inside of a circle of radius $1/2$ centered at $(1, 0)$ in the complex plane (and similarly for Y). Hence for $|t| < \frac{1}{C_7}$ the $\ln \mathbf{E}[\exp(itX)]$ is holomorphic, where ln denotes the principal branch of the logarithm. Since $|\ln u - \ln v| \leqslant 2\,|u-v|$ for any complex number u and v with $|u-1| \leqslant \frac{1}{2}$ and $|v-1| \leqslant \frac{1}{2}$, for $|t| < \frac{1}{C_7}$ we have

$$\begin{aligned}
&|\ln \mathbf{E}[\exp(itX)] - \ln \mathbf{E}[\exp(itY)]| \\
&\quad \leqslant 2\,|\mathbf{E}[\exp(itX)] - \mathbf{E}[\exp(itY)]| = 2\,|\mathbf{E}[\exp(itX)(1 - \exp(it(Y-X)))]| \\
&\quad \leqslant 2\mathbf{E}[|\exp(itX) - 1|\,|1 - \exp(it(Y-X))|] + 2\mathbf{E}[|1 - \exp(it(Y-X))|] \\
&\quad \leqslant 2t^2 \mathbf{E}[|X|\,|Y-X|] + t^2 \mathbf{E}[(Y-X)^2] \\
&\quad \leqslant 4t^2 C_7 \sqrt{\mathbf{E}[(X-Y)^2]},
\end{aligned}$$

where in the last part we used Cauchy–Schwarz inequality and the triangle inequality. The second part of Lemma 7 is standard and the proof can be found in many basic probability textbooks (see, for example, ref. 2, p. 103).

Proof of Proposition 5. Let ϕ_n denote the characteristic function of the normalized random variable $\tilde{x}_n$:

$$\phi_n(t) = \mathbf{E}[\exp(it\tilde{x}_n)]$$

We shall prove that there exists a positive constant C_8 such that for any $|t| < \frac{1}{C_8}$

$$\lim_{n \to \infty} \ln \phi_n(t) = -\tfrac{1}{2} \sigma_\infty^2 t^2.$$

By continuity theorem (see ref. 2, p. 99) this will imply that $\tilde{x}_n$ converges in distribution to a normal variable with mean zero and variance σ_∞^2 and, in view of (9) the proof will be complete.

From Lemmas 6 and 7, it follows that there exists a positive constant C_8 such that for any $|t| < \frac{1}{C_8}$ and any $n \geqslant 1$

$$\left|\ln \mathbf{E}[\exp(it\tilde{x}_{n+1})] - \ln \mathbf{E}\left[\exp\left(it\frac{1}{\lambda_n}\tilde{y}_n\right)\right]\right| \leqslant 4C_8 t^2 \sqrt{\mathbf{Var}\left[\frac{\tilde{z}_n}{\lambda_n}\right]}.$$

For all n and $i \leqslant k$, let $\beta_{n,i} = \frac{\alpha_{n,i}}{\lambda_n}$. Then from (7) and (8), there is a positive constant C_9 such that for all $|t| < \frac{1}{C_8}$ and any $n \geqslant 1$

$$\left|\ln \phi_{n+1}(t) - \sum_{i=1}^{k} \ln \phi_n(\beta_{n,i}t)\right| \leqslant C_9 t^2 (1-\delta_2)^n.$$

Now, let $\gamma_{n,j}$ be any real numbers satisfying $\sum_j \gamma_{n,j}^2 = 1$. Applying the above inequality with $\gamma_{n,j}t$ in place of t and summing over j we get

$$\sum_j \left|\ln \phi_{n+1}(\gamma_{n,j}t) - \sum_{i=1}^{k} \ln \phi_n(\gamma_{n,j}\beta_{n,i}t)\right| \leqslant C_9 t^2 (1-\delta_2)^n.$$

It follows that for all $|t| < \frac{1}{C_8}$

$$\begin{aligned} &\left|\ln \phi_{n+m}(t) - \sum_{i_1, i_2, \ldots, i_m} \ln \phi_n(t\beta_{n,i_1}\beta_{n+1,i_2}\cdots\beta_{n+m-1,i_m})\right| \\ &\qquad \leqslant C_9 t^2 \sum_{p=0}^{m-1} (1-\delta_2)^{n+p} \leqslant \frac{C_9}{\delta_2} t^2 (1-\delta_2)^n. \end{aligned} \tag{10}$$

Now, by assumption, at least two α_i are nonzero which implies that

$$\max_{i_1, i_2, \ldots, i_m} \beta_{n,i_1}\beta_{n+1,i_2}\cdots\beta_{n+m-1,i_m}$$

converges to zero when $m \to \infty$. From the second part of Lemma 7, it follows that for any n and any $\epsilon_1 > 0$ there exists an $M_1(n, \epsilon_1)$ such that for all $m \geqslant M_1$

$$\begin{aligned} &|\ln \phi_n(t\beta_{n,i_1}\beta_{n+1,i_2}\cdots\beta_{n+m-1,i_m}) + \tfrac{1}{2}t^2(\beta_{n,i_1}\beta_{n+1,i_2}\cdots\beta_{n+m-1,i_m})^2\,\mathbf{Var}[\tilde{x}_n]| \\ &\qquad \leqslant \epsilon_1 t^2 (\beta_{n,i_1}\beta_{n+1,i_2}\cdots\beta_{n+m-1,i_m})^2. \end{aligned}$$

Hence for any $m \geqslant M_1$, by summing the above inequality over all $i_1, i_2,..., i_m$, we obtain

$$\left| \sum_{i_1, i_2,..., i_m} \ln \phi_n(t\beta_{n, i_1}\beta_{n+1, i_2} \cdots \beta_{n+m-1, i_m}) + \tfrac{1}{2}t^2 \mathbf{Var}[\tilde{x}_n] \right| \leqslant \epsilon_1 t^2. \qquad (11)$$

From (10) and (11), using the triangle inequality, we have for all n, $\epsilon_1 > 0$, and $|t| < \frac{1}{C_8}$

$$\limsup_{m \to \infty} \left| \ln \phi_{n+m}(t) + \frac{1}{2} t^2 \mathbf{Var}[\tilde{x}_n] \right| \leqslant \epsilon_1 t^2 + \frac{C_9}{\delta_2} t^2 (1-\delta_2)^n.$$

Since $\epsilon_1 > 0$ is arbitrary, the proof is finished by taking the limit $n \to \infty$.

4. PROOF OF THE MAIN THEOREM

In order to apply Proposition 5 to a hierarchical sequence, we need to verify that the sequence satisfies the required variance bounds. As a first step in proving the upper bound we will use the following estimate of the large deviations of x_n. While this estimate is probably not optimal (see the remark following the statement of the lemma), it is sufficient to prove the variance bounds necessary to verify the assumptions of Proposition 5.

Lemma 8. Let a and b be real numbers with $a < b$, and $f: [a, b]^k \to [a, b]$ be averaging. Also let x_0 be an $[a, b]$-valued random variable. Furthermore assume that there exists a $c \in [a, b]$ such that x_n, defined as in (2), converges to c in probability. Then for any $\epsilon > 0$ there exists an $M > 0$ such that for all n

$$\mathbf{P}[|x_n - c| > \epsilon] \leqslant M\epsilon^n$$

Remark. Roughly speaking, the lemma says that $\mathbf{P}[|x_n - c| > \epsilon]$ decays faster than ϵ^n for any ϵ. In fact, based on analogy with large deviation probabilities for sums of independent, identically distributed random variables (ref. 2, p. 74), we would expect an upper bound of the form $\exp[-I(\epsilon)\, C^n]$ with a $C > 1$. It would be interesting to prove such a bound for hierarchical sequences considered in this paper.

Proof of this lemma is given in the Appendix.

From now on, until the end of the section, we assume all the assumptions of Theorem 2. Let us remind that c denotes the limit of x_n in probability.

Let also $c_n = \mathbf{E}[x_n]$. Then, since the x_n are uniformly bounded, we also have $c = \lim_{n\to\infty} c_n$. Let us define

$$\begin{aligned} \alpha_{n,i} &= \frac{\partial f}{\partial u_i}(c_n, c_n, \ldots, c_n) \\ \alpha_i &= \frac{\partial f}{\partial u_i}(c, c, \ldots, c) \\ \lambda_n &= \sqrt{\sum_{i=1}^{k} \alpha_{n,i}^2} \\ \lambda &= \sqrt{\sum_{i=1}^{k} \alpha_i^2} \\ \underline{\lambda} &= \inf_n \lambda_n. \end{aligned} \tag{12}$$

By expanding f at $(c_n, c_n, \ldots, c_n)$, we get

$$x_{n+1} = f(c_n, c_n, \ldots, c_n) + \sum_{i=1}^{k} \alpha_{n,i}(x_{n,i} - c_n) + O\left(\sum_{i=1}^{k} (x_{n,i} - c_n)^2\right).$$

For each n let

$$d_n = f(c_n, c_n, \ldots, c_n) - c_n \sum_{i=1}^{k} \alpha_{n,i} \tag{13}$$

and let

$$z_n = x_{n+1} - \sum_{i=1}^{k} \alpha_{n,i} x_{n,i}. \tag{14}$$

Then

$$z_n = d_n + O\left(\sum_{i=1}^{k} (x_{n,i} - c_n)^2\right). \tag{15}$$

Note that (14) can be viewed as a relation in Proposition 5. Hence in order to prove the main theorem we only need to show that x_n and z_n (defined as in (2) and (14)) satisfy the variance bounds in Proposition 5. Note that by property 1 in Definition 1, for every u we have $f(u, u, \ldots, u) = u$; differentiating this relation at $u = c$ gives have $\sum \alpha_i = 1$ and, since $\alpha_i \geqslant 0$ and least two of the α_i are different from zero, this implies $0 < \lambda < 1$.

We now use Lemma 8 to prove Propositions 9 and 10. Theorem 2 will then follow from Proposition 5. To prove Propositions 9 and 10, we need the following additional notation: Let

$$y_n = \sum_{i=1}^{k} \alpha_{n,i} x_{n,i}. \tag{16}$$

Then

$$x_{n+1} = y_n + z_n \tag{17}$$

Let the family $\{x'_{n,i}: n=0, 1, 2,..., i=1, 2,..., k\}$ be an independent copy of $\{x_{n,i}: n=0, 1, 2,..., i=1,..., k\}$, i.e., the families are independent and the joint distribution of the variables in the first family is the same as that of the variables in the second. Let

$$x'_n = x'_{n,1}.$$

The sequence x'_n is then an independent copy of the sequence x_n. In the way analogous to (16) and (17) we define

$$y'_n = \sum_{i=1}^{k} \alpha_{n,i} x'_{n,i}.$$

and

$$z'_n = x'_{n+1} - y'_n.$$

We have the following variance bounds.

Proposition 9. For any $\epsilon > 0$ there exists an $M > 0$ such that for all n

$$\mathbf{Var}[z_n] \leqslant M(\lambda + \epsilon)^{4n}.$$

Proof. Since f is twice continuously differentiable in the neighborhood of $(c, c,..., c)$, it follows from (15) that

$$\mathbf{Var}[z_n] \leqslant M_1 \mathbf{E}[(x_n - c_n)^4] \tag{18}$$

with an $M_1 > 0$. Writing

$$x_n - x'_n = (x_n - c_n) - (x'_n - c_n),$$

expanding and using the independence of x_n and x'_n we obtain

$$\mathbf{E}[(x_n - x'_n)^4] = 2\mathbf{E}[(x_n - c_n)^4] + 6(\mathbf{E}[(x_n - c_n)^2])^2,$$

whence, using (18) we get

$$\mathbf{Var}[z_n] \leqslant \frac{M_1}{2} \mathbf{E}[(x_n - x'_n)^4]. \tag{19}$$

To estimate the right hand side of the above inequality, we proceed inductively. First note that

$$\begin{aligned}\mathbf{E}[(y_n - y'_n)^4] &\leqslant \left(\sum_{i=1}^{k} \alpha_{n,i}^4\right) \mathbf{E}[(x_n - x'_n)^4] + \lambda_n^4(\mathbf{E}[(x_n - x'_n)^2])^2 \\ &\leqslant \lambda_n^4(\mathbf{E}[(x_n - x'_n)^4] + (\mathbf{E}[(x_n - x'_n)^2])^2).\end{aligned}$$

Since for all real a and b, $(a+b)^4 \leqslant 8(a^4 + b^4)$ we have

$$\mathbf{E}[(z_n - z'_n)^4] \leqslant 8(\mathbf{E}[(z_n - d_n)^4] + \mathbf{E}[(z'_n - d_n)^4]).$$

Hence from (15) there exists an $M > 0$ such that for all n

$$\mathbf{E}[(z_n - z'_n)^4] \leqslant M\mathbf{E}[(x_n - c_n)^8].$$

Since for any $\epsilon > 0$ there exists M such that for all a and b $(a+b)^4 \leqslant (1+\epsilon)\, a^4 + Mb^4$, for such ϵ there exists M' such that

$$\begin{aligned}&\mathbf{E}[(x_{n+1} - x'_{n+1})^4] \\ &\quad = \mathbf{E}[(y_n - y'_n + z_n - z'_n)^4] \\ &\quad \leqslant (1+\epsilon)\, \mathbf{E}[(y_n - y'_n)^4] + M\mathbf{E}[(z_n - z'_n)^4] \\ &\quad \leqslant (1+\epsilon)\, \lambda_n^4(\mathbf{E}[(x_n - x'_n)^4] + (\mathbf{E}[(x_n - x'_n)^2])^2) + M'\mathbf{E}[(x_n - c_n)^8].\end{aligned}$$

Since, by the large deviation estimate of Lemma 8, for any $\epsilon > 0$ there exists N such that for all $n \geqslant N$

$$M'E[(x_n - c_n)^8] \leqslant \epsilon E[(x_n - c_n)^4] + \epsilon^n.$$

It follows that for any ϵ there exists an N such that for all $n > N$

$$\mathbf{E}[(x_{n+1} - x'_{n+1})^4] \leqslant (\lambda + \epsilon)^4\, (\mathbf{E}[(x_n - x'_n)^4] + \mathbf{E}[(x_n - x'_n)^2]^2) + \epsilon^n. \tag{20}$$

Since

$$\mathbf{E}[(x_n - x_n')^2] = 2\,\mathbf{Var}[x_n],$$

in order to estimate the second term in (20), it is sufficient to estimate the variance of x_n. By the triangle inequality we have

$$\sqrt{\mathbf{Var}[x_{n+1}]} \leqslant \lambda_n \sqrt{\mathbf{Var}[x_n]} + \sqrt{\mathbf{Var}[z_n]}.$$

For any $\epsilon_1 > 0$, from (18) we have

$$\mathbf{Var}[z_n] \leqslant M_1 \epsilon_1^2 \mathbf{E}[(x_n - c_n)^2;\, |x_n - c_n| \leqslant \epsilon_1] + M_1 \mathbf{E}[(x_n - c_n)^4;\, |x_n - c_n| \geqslant \epsilon_1].$$

Applying the large deviation estimate of Lemma 8 to the above inequality, we prove that for any $\epsilon > 0$ there exists $M > 0$ such that for all n

$$\mathbf{Var}[z_n] \leqslant \epsilon^2\,\mathbf{Var}[x_n] + M^2 \epsilon^{2n}.$$

It now follows from the triangle inequality that

$$\sqrt{\mathbf{Var}[z_n]} \leqslant \epsilon \sqrt{\mathbf{Var}[x_n]} + M\epsilon^n.$$

Hence

$$\sqrt{\mathbf{Var}[x_{n+1}]} \leqslant (\lambda_n + \epsilon) \sqrt{\mathbf{Var}[x_n]} + M\epsilon^n$$

and it follows by induction on n that for any $\epsilon > 0$ there exists an $M' > 0$ such that

$$\mathbf{E}[(x_n - x_n')^2] \leqslant M'(\lambda + \epsilon)^{2n}. \tag{21}$$

From (20) and (21) it easily follows that for any $\epsilon > 0$ there exist M, N such that for all $n > N$

$$\mathbf{E}[(x_{n+1} - x_{n+1}')^4] \leqslant (\lambda + \epsilon)^4\,\mathbf{E}[(x_n - x_n')^4] + M(\lambda + \epsilon)^{4n} + \epsilon^n.$$

By induction on n it now follows that for any $\epsilon > 0$ there exist M, N such that for all $n > N$

$$\mathbf{E}[(x_n - x_n')^4] \leqslant M(\lambda + \epsilon)^{4n}.$$

Together with (19) this proves the proposition.

Proposition 10. For any $\epsilon > 0$ there exists an $M > 0$ such that for all n

$$\mathbf{Var}[x_n] > M(\lambda - \epsilon)^{2n}$$

Proof. For all $\epsilon_1 > 0$ and n let

$$C(n, \epsilon_1) = (\mathbf{E}[(x_n - x'_n)^2; |x_n - c| \leqslant \epsilon_1; |x'_n - c| \leqslant \epsilon_1])^{\frac{1}{2}}.$$

Since $C(n, \epsilon_1)^2 \leqslant 2\,\mathbf{Var}[x_n]$ it suffices to show that for any $\epsilon > 0$ there exist $\epsilon_1 > 0$, $M > 0$ such that

$$C(n, \epsilon_1) > M(\lambda - \epsilon)^n.$$

Take any $\epsilon > 0$. By the mean value theorem, there exists an $\epsilon_1 > 0$ such that for all $|x_{n,i} - c| \leqslant \epsilon_1$ and $|x'_{n,i} - c| \leqslant \epsilon_1$ where $i = 1, 2, \ldots, k$, we have

$$\left| x_{n+1} - x'_{n+1} - \sum_{i=1}^{k} \alpha_i (x_{n,i} - x'_{n,i}) \right| \leqslant \epsilon \sum_{i=1}^{k} |x_{n,i} - x'_{n,i}|. \tag{22}$$

For such $x_{n,i}$, $x'_{n,i}$ we have

$$|x_{n+1} - c| \leqslant \epsilon_1 \qquad \text{and} \qquad |x'_{n+1} - c| \leqslant \epsilon_1.$$

Hence

$$C(n+1, \epsilon_1) \geqslant (\mathbf{E}[(x_{n+1} - x'_{n+1})^2; |x_{n,i} - c| \leqslant \epsilon_1, |x'_{n,i} - c| \leqslant \epsilon_1, \text{for } i = 1, \ldots, k])^{\frac{1}{2}}$$

and so, by the triangle inequality and (22) $C(n+1, \epsilon_1)$ is bigger than or equal to

$$\left(\mathbf{E}\left[\left(\sum_{i=1}^{k} \alpha_i (x_{n,i} - x'_{n,i})\right)^2; |x_{n,i} - c| \leqslant \epsilon_1, |x'_{n,i} - c| \leqslant \epsilon_1, \text{for } i = 1, \ldots, k\right]^{\frac{1}{2}}\right.$$
$$\left. - \epsilon\left(\mathbf{E}\left[\left(\sum_{i=1}^{k} |(x_{n,i} - x'_{n,i})|\right)^2; |x_{n,i} - c| \leqslant \epsilon_1, |x'_{n,i} - c| \leqslant \epsilon_1, \text{for } i = 1, \ldots, k\right]\right)^{\frac{1}{2}}.$$

Hence using the independence of the $x_{n,i} - x'_{n,i}$ for different i in the first term and applying the triangle inequality to the second term we get

$$C(n+1, \epsilon_1) \geqslant \lambda \mathbf{P}[|x_n - c| \leqslant \epsilon_1]^{k-1} C(n, \epsilon_1) - \epsilon k C(n, \epsilon_1).$$

By the weak law of large numbers, there exists an $N > 0$ such that for all $n \geqslant N$,

$$C(n+1, \epsilon_1) \geqslant (\lambda(1-\epsilon) - k\epsilon)\, C(n, \epsilon_1).$$

The last step of the proof is to show that there exists an $N_1 > N$ such that

$$C(N_1, \epsilon_1) > 0,$$

since it then follows by induction that for any ϵ there exist $M_1 > 0$, such that for all $n \geqslant N_1$

$$C(n, \epsilon_1) > M_1(\lambda(1-\epsilon) - k\epsilon)^n$$

and this proves the proposition.

To prove existence of such an N_1, it suffices to show that for any $\epsilon > 0$ there exists an N such that for all $n \geqslant N$ the range of x_n contains at least two distinct points in the interval $(c-\epsilon, c+\epsilon)$. Denote the range of x_n by

$$\mathscr{R}(x_n) = \{u \in [a, b] : \text{ for any } \epsilon > 0 \;\; \mathbf{P}[|x_n - u| < \epsilon] > 0\}.$$

Since the range of x_0 consists of more than one point, we can without loss of generality assume that there exist u and v in $\mathscr{R}(x_0)$ such that $u \leqslant c < v$. Let $c_n^* = \inf\{y > c\colon y \in \mathscr{R}(x_n)\}$. Then $c_n^* \in \mathscr{R}(x_n)$. Since $\mathscr{R}(x_{n+1}) \supset \mathscr{R}(x_n)$ (by averaging property of f) c_n^* is nonincreasing and bounded below by c. Let $c^* = \lim c_n^*$. We will prove $c^* = c$ by contradiction. Since x_n converges to c in probability, we can choose $w_n \in \mathscr{R}(x_n)$ such that w_n converges to c. Suppose that $c^* > c$. Then by averaging property of f there exists $u_i \in \{c, c^*\}$ such that

$$c < f(u_1, u_2, \ldots, u_k) < c^*.$$

For each n and i let

$$u_{n,i} = \begin{cases} w_n & \text{if} \quad u_i = c \\ c_n^* & \text{if} \quad u_i = c^* \end{cases}$$

Then, by the continuity of f, for large n we have

$$c < f(u_{n,1}, u_{n,2}, \ldots, u_{n,k}) < c^*$$

Since $u_{n,i} \in \mathscr{R}(x_n)$, we get

$$f(u_{n,1}, u_{n,2},..., u_{n,k}) < c^* \leqslant c^*_{n+1},$$

which is a contradiction. Therefore $c^* = c$. Hence either $c^*_n = c$ for some n or $c^*_n > c$ and c^*_n converges to c. In either case, it easily follows that for any ϵ there exists an N such that for all $n \geqslant N$ the range of x_n contains at least two points in the interval $(c, c+\epsilon)$ and the proof is complete.

APPENDIX

Proof of Proposition 3. For the ith bond of **G** let $s_1(i)$ and $s_2(i)$ be its two end sites. An equivalent expression for f is given by the Dirichlet variational principle:

$$f(u_1, u_2,..., u_k) = C \min_V \sum_{i=1}^{k} u_i[V(s_1(i)) - V(s_2(i))]^2 \tag{23}$$

where the minimum is taken over all real-valued functions V defined on the site set of **G** with $V(s_t) = 1$ and $V(s_b) = 0$ and the normalization constant $C > 0$ is chosen to satisfy $f(1, 1,..., 1) = 1$ (see ref. 1 for the variational principle). Assumption 2 of Definition 1 is called the Rayleigh's monotonicity law. We will verify it here for completeness using the variational principle (see ref. 1 for another proof). Take $u_1, u_2,..., u_k, u'_1, u'_2,..., u'_k$ with $u_i \leqslant u'_i$ where $i = 1, 2,..., k$. Let V^* be a non-negative real-valued function defined on the site set of **G** with $V^*(s_t) = 1$ and $V^*(s_b) = 0$ which minimizes (23) with the values of the bond conductivities $(u'_1, u'_2,..., u'_k)$ (this V^* is called the potential). Since

$$f(u_1, u_2,..., u_k) \leqslant C \sum_{i=1}^{k} u_i[V^*(s_1(i)) - V^*(s_2(i))]^2 \leqslant f(u'_1, u'_2,..., u'_k),$$

Assumption 2 follows. To verify assumption 1 it is now enough to note that $f(u, u,..., u) = uf(1, 1,..., 1) = u$ (recall that the normalization constant C was chosen to satisfy $f(1, 1,..., 1) = 1$). Now we prove Assumption 3. Take two nonnegative numbers $u < v$ and two distinct indices $i_1, i_2 \in \{1, 2,..., k\}$. Let $u_{i_1} = u$ and $u_{i_2} = v$. Let V^* be a function defined on the site set of **G** which minimizes (23) in the case when all the conductivities are equal to v. By homogeneity, the same function V^* also minimizes (23) at $(u, u,..., u)$. Since $V^*(s_t) = 1, V^*(s_b) = 0$, there exist at least two distinct bonds such that at the endpoints of each one of these bonds V^* takes different values. This is true because we have at least two disjoint self-avoiding paths connecting the top site to the bottom site and one such bond can be chosen from each one of these two paths. We can assume without loss of generality that b_1, b_2 are such bonds and that the index $i_2 \neq 1$. Let $u_1 = u$, $u_{i_1} = u$ and let $u_i = v$ for all other indices $i \in \{2,..., k\}$. Since $u_1 < v$ and V^* has different values at the end points of the bond b_1, from the variational principle we have

$$f(u_1, u_2,..., u_k) < f(v, v,..., v) = v.$$

To prove the other inequality of Assumption 3 for the same choice of u_i, we use the dual variational principle (called Thomson's principle) for resistivity (see ref. 1). Recall that resistivity is defined as the reciprocal of conductivity.Thomson's principle is (see ref. 1):

$$\frac{1}{f(u_1, u_2,..., u_k)} = C^* \min_I \sum_{i=1}^{k} \frac{1}{u_i} I(b_i)^2. \tag{24}$$

In the above sum, the minimum is taken over all real-valued functions I defined on the bond set of $\mathbf{G}$ satisfying the constraint

$$\sum_{b \sim s} I(b) = \begin{cases} 0 & \text{if } s \text{ is a internal site} \\ 1 & \text{if } s \text{ is the top site} \\ -1 & \text{if } s \text{ is the bottom site} \end{cases}$$

where $b \sim s$ means that the sum is taken over all bonds b adjacent to the site s and the C^* in the above sum is chosen to satisfy $f(1, 1,..., 1) = 1$. Using Thomson's principle we obtain, by an argument similar to the above the following inequality for resistivity

$$\frac{1}{f(u_1, u_2,..., u_k)} < \frac{1}{u}.$$

This is true even in the case $u = 0$, since the right side is then infinite while the left side is still finite, because there exists a path consisting of resistors with finite resistivities ($\frac{1}{v} < \infty$) connecting the surface sites.

Proof of Lemma 8. Let

$$Q = \{u \in [a, c) : \text{for any } \epsilon > 0 \text{ there exists an } M \text{ such that}$$
$$\text{for all } n \ \mathbf{P}[x_n < u] \leqslant M\epsilon^n\}.$$

Clearly, Q is an interval. First we will prove that $\sup Q = c$. Since $\mathbf{P}[x_n < a] = 0$, $a \in Q$. Suppose $\sup Q < c$. First by Assumption 3 of Definition 1,

$$f(u_1, u_2, \ldots, u_k) > \sup Q$$

for $u_{\sigma(1)} = \sup Q$, $u_{\sigma(2)} = c, \ldots, u_{\sigma(k)} = c$ and any permutation σ of $\{1, 2, \ldots, k\}$. By continuity of f we can choose an $\epsilon \in (0, c)$ such that

$$f(u_1, u_2, \ldots, u_k) > \sup Q + \epsilon \tag{25}$$

and

$$u_{\sigma(1)} = \sup Q - \epsilon,\ u_{\sigma(2)} = c - \epsilon, \ldots, u_{\sigma(k)} = c - \epsilon.$$

Suppose $f(u_1, u_2, \ldots, u_k) < \sup Q + \epsilon$. Choose a permutation σ of $\{1, 2, \ldots, k\}$ such that $u_{\sigma(1)} \leqslant u_{\sigma(2)} \leqslant \cdots \leqslant u_{\sigma(k)}$. Then by the Assumption 1 of Definition 1 $u_{\sigma(1)} < \sup Q + \epsilon$. If $u_{\sigma(1)} \geqslant \sup Q - \epsilon$ then by (25) $u_{\sigma(2)}$ cannot be bigger than or equal to $c - \epsilon$. Hence we have either $u_{\sigma(1)} < \sup Q - \epsilon$ or $u_{\sigma(1)} < \sup Q + \epsilon$ and $u_{\sigma(2)} < c - \epsilon$. Hence

$$\begin{aligned}
&\mathbf{P}[x_{n+1} < \sup Q + \epsilon] \\
&\quad \leqslant \sum_{\sigma} \mathbf{P}[f(x_{n,1}, x_{n,2}, \ldots, x_{n,k}) < \sup Q + \epsilon; \\
&\qquad\quad x_{n,\sigma(1)} \leqslant x_{n,\sigma(2)} \leqslant \cdots \leqslant x_{n,\sigma(k)};\ x_{n,\sigma(1)} < \sup Q + \epsilon] \\
&\quad \leqslant \sum_{\sigma} \mathbf{P}[x_{n,\sigma(1)} < \sup Q - \epsilon] + \sum_{\sigma} \mathbf{P}[x_{n,\sigma(1)} < \sup Q + \epsilon;\ x_{n,\sigma(2)} < c - \epsilon],
\end{aligned}$$

where the sum is taken over all permutation σ of $\{1, 2, \ldots, k\}$. Hence using the independence of $x_{n,\sigma(1)}$ and $x_{n,\sigma(2)}$, we have

$$\begin{aligned}
&\mathbf{P}[x_{n+1} < \sup Q + \epsilon] \\
&\qquad \leqslant k!\, \mathbf{P}[x_n < \sup Q - \epsilon] + k!\, \mathbf{P}[x_n < c - \epsilon]\, \mathbf{P}[x_n < \sup Q + \epsilon].
\end{aligned}$$

Take any $\epsilon_1 > 0$. By definition of Q and by the weak law of the large numbers, the above inequality yields existence of N and M such that for all $n \geqslant N$

$$\mathbf{P}[x_{n+1} < \sup Q + \epsilon] < M \left(\frac{\epsilon_1}{2}\right)^n + \frac{\epsilon_1}{2} \mathbf{P}[x_n < \sup Q + \epsilon]$$

From this it follows by an easy inductive argument that there exists an M' (depending on ϵ_1) such that

$$\mathbf{P}[x_n < \sup Q + \epsilon] < M'\epsilon_1^n.$$

Since $\epsilon_1 > 0$ is arbitrary it follows that $\sup Q + \epsilon \in Q$, which is a contradiction.

We have shown that for any $\epsilon > 0$ and for any $u < c$ there exists $M > 0$ such that for all n

$$\mathbf{P}[x_n < u] \leqslant M\epsilon^n.$$

Applying this result to the averaging function $g(x) = -f(-x)$ we prove that for any $\epsilon > 0$ and for any $u > c$ there exists an $M > 0$ such that for all n

$$\mathbf{P}[x_n > u] \leqslant M\epsilon^n.$$

Proof of the lemma is finished.

ACKNOWLEDGMENT

Both authors were partially supported by the NSF Statistics and Probability Grant DMS-9706915.

REFERENCES

1. P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks* (Math. Assoc. Amer., Washington DC, 1984).
2. R. Durrett, *Probability: Theory and Examples*, 2nd ed. (Duxbury Press, 1996).
3. R. B. Griffiths and M. Kaufman, Spin systems on hierarchical lattices. Introduction and thermodynamic limit, *Phys. Rev. B*.**26**:5022–5032 (1982).
4. Deli Li and T. D. Rogers, Asymptotic behavior for iterated functions of random variables, *Ann. Appl. Probab.* **9**(4):1175–1201 (1999).
5. A. Schenkel, J. Wehr, and P. Wittwer, Computer assisted proofs for fixed-point problems in Sobolev spaces, *Mathematical Physics Electronic Journal* **6**(3) (2000) (at http://mpej.unige.ch/mpej/MPEJ.html).
6. T. Schlösser and H. Spohn, Sample-to-sample fluctuations in the conductivity of a disordered medium, *J. Stat. Phys.* **69**:955–967 (1992).
7. I. Shneiberg, Hierarchical sequences of random variables, *Prob. Theor. Appl.* **31**:137–141 (1986).