Elect. Comm. in Probab. **4** (1999) 39–49

ELECTRONIC COMMUNICATIONS in PROBABILITY

TREES, NOT CUBES: HYPERCONTRACTIVITY, COSINESS, AND NOISE STABILITY

Oded SCHRAMM
The Weizmann Institute of Science, Rehovot 76100, Israel
schramm@wisdom.weizmann.ac.il
http://www.wisdom.weizmann.ac.il/∼schramm/

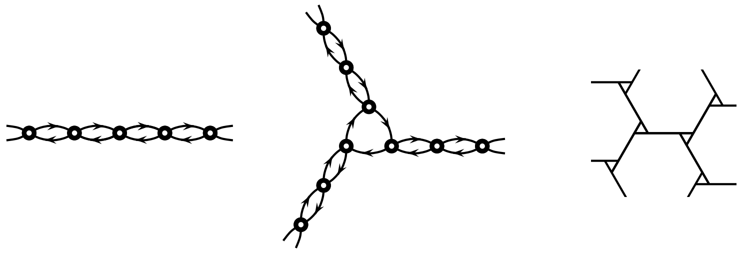
Boris TSIRELSON
School of Mathematics, Tel Aviv Univ., Tel Aviv 69978, Israel
tsirel@math.tau.ac.il
http://math.tau.ac.il/∼tsirel/

submitted February 25, 1999, *revised* July 22, 1999

Abstract:
Noise sensitivity of functions on the leaves of a binary tree is studied, and a hypercontractive inequality is obtained. We deduce that the spider walk is not noise stable.

Introduction

For the simplest random walk (Fig. 1a), the set Ω_n^{simp} of all n-step trajectories may be thought of either as (the set of leaves of) a binary tree, or (the vertices of) a binary cube $\{-1,+1\}^n$. However, consider another random walk (Fig. 1b); call it the simplest *spider* walk, since it is a discrete counterpart of a spider martingale, see [2]. The corresponding Ω_n^{spider} is the set of leaves of a binary tree. It is not quite appropriate to think of such n-step "spider walks" as the vertices of a binary cube, since for different i and j in $\{1,2,\dots,n\}$ it is not necessary that the j'th step has the same or opposite direction from the i'th step. Of course, one may choose to ignore this point, and use the n bits given by a point in $\{-1,1\}^n$ to describe a spider walk, in such a way that for each $j=1,2,\dots,n$, the first j bits determine the first j steps of the walk. Such a correspondence would not be unique. To put it differently, the vertices of the cube have a natural associated partial order. When you consider two walks on $\mathbb{Z}$, the associated partial ordering has a natural interpretation, one trajectory is (weakly) larger than the other if whenever the latter moved to the right, the former also moved to the right. However, this interpretation does not make sense for the spider walk in Fig. 1b, and even less to more complicated "spider webs" with several "roundabouts", such as that of Fig. 1c.
Noise sensitivity and stability are introduced and studied in [3] for functions on cubes. Different cube structures on a binary tree are non-equivalent in that respect. It is shown here that a



(a) (b) (c)

Figure 1: (a) simple walk; (b) spider walk; (c) a spider web. At each point, there are two equiprobable moves.

natural function on Ω_n^{spider} is non-stable under *every* cube structure. One of the tools used is a new hypercontractive inequality, which hopefully may find uses elsewhere.

1 Stability and sensitivity on cubes, revisited

A function $f : \{-1,+1\}^n \to \mathbb{C}$ has its Fourier-Walsh expansion,

$$\begin{aligned} f(\tau_1,\dots,\tau_n) &= \\ &= \hat f_0 + \sum_k \hat f_1(k)\tau_k + \sum_{k<l} \hat f_2(k,l)\tau_k\tau_l + \dots + \hat f_n(1,\dots,n)\tau_1\dots\tau_n\,. \end{aligned}$$

Set

$$\tilde f_j(\tau_1,\dots,\tau_n) = \sum_{i_1<i_2<\dots<i_j} \hat f_j(i_1,\dots,i_j)\tau_{i_1}\tau_{i_2}\dots\tau_{i_j}\,.$$

Since the transform $f \mapsto \hat f$ is isometric, we have $\|f\|^2 = \sum_0^n \|\tilde f_m\|^2$, where

$$\|f\|^2 = 2^{-n} \sum_{\tau_1,\dots,\tau_n} |f(\tau_1,\dots,\tau_n)|^2\,. \tag{1.1}$$

The quantities

$$S_1^m(f) = \sum_{i=1}^m \|\tilde f_i\|^2\,, \quad S_m^\infty(f) = \sum_{i=m}^n \|\tilde f_i\|^2$$

are used for describing low-frequency and high-frequency parts of the spectrum of f.
Given a sequence of functions $F = \big(f_n\big)_{n=1}^\infty$, $f_n : \{-1,+1\}^n \to \mathbb{C}$, satisfying $0 < \liminf_{n\to\infty} \|f_n\| \le \limsup_{n\to\infty} \|f_n\| < \infty$, we consider numbers

$$\begin{aligned} S_1^m(F) &= \limsup_{n\to\infty} S_1^m(f_n)\,, \\ S_m^\infty(F) &= \limsup_{n\to\infty} S_m^\infty(f_n)\,. \end{aligned}$$

Here is one of equivalent definitions of stability and sensitivity for such F, according to [3, Th. 1.8] (indicator functions are considered there):

$$\begin{aligned} F \text{ is stable} \quad &\text{iff} \quad S^\infty_m(F) \to 0 \text{ for } m \to \infty\,, \\ F \text{ is sensitive} \quad &\text{iff} \quad S^m_1(F) = 0 \text{ for all } m\,. \end{aligned}$$

Constant components are irrelevant; that is, if $g_n = f_n + c_n$, $c_n \in \mathbb{C}$, then $S^m_1(f_n) = S^m_1(g_m)$ and $S^\infty_{m+1}(f_n) = S^\infty_{m+1}(g_n)$, therefore stability of $\big(f_n\big)_{n=1}^\infty$ is equivalent to stability of $\big(g_n\big)_{n=1}^\infty$; the same for sensitivity. If $\big(f_n\big)_{n=1}^\infty$ is both stable and sensitive, then (and only then) $S^\infty_1(f_n) \to 0$, that is, $\|f_n - c_n\| \to 0$ for some $c_n \in \mathbb{C}$.
A random variable τ will be called a random sign, if $\mathbb{P}(\tau = -1) = 1/2$ and $\mathbb{P}(\tau = +1) = 1/2$. A joint distribution for two random signs τ', τ'' is determined by their correlation coefficient $\rho = \mathbb{E}(\tau'\tau'') = 1 - 2\mathbb{P}(\tau' \ne \tau'')$. Given n independent pairs $(\tau'_1, \tau''_1), \dots, (\tau'_n, \tau''_n)$ of random signs with the same correlation ρ for each pair, we call $(\tau'_1, \dots, \tau'_n)$ and $(\tau''_1, \dots, \tau''_n)$ a *ρ-correlated pair of random points of the cube* $\{-1,+1\}^n$. (In terms of [3] it is $\big(x, N_\varepsilon(x)\big)$ with $\varepsilon = (1-\rho)/2$.) It is easy to see that

$$\mathbb{E}\big(\overline{f(\tau')} f(\tau'')\big) = \sum_{m=0}^n \rho^m \|\tilde f_m\|^2$$

for a ρ-correlated pair (τ', τ''). We may write it as a scalar product in the space $L_2\big(\{-1,+1\}^n\big)$ with the norm (1.1),

$$\mathbb{E}\big(\overline{f(\tau')} f(\tau'')\big) = (\rho^{\mathbf{N}} f, f)\,; \tag{1.2}$$

here $\rho^{\mathbf{N}}$ is the operator $\rho^{\mathbf{N}} f = \sum_n \rho^n \tilde f_n$. Similarly, $\mathbb{E}\big(\overline{g(\tau')} f(\tau'')\big) = (\rho^{\mathbf{N}} f, g)$. On the other hand,

$$\mathbb{E}\big(\overline{g(\tau')} f(\tau'')\big) = \mathbb{E}\big(\overline{g(\tau')} \cdot \mathbb{E}\big(f(\tau'')|\tau'\big)\big) = \big(\tau' \mapsto \mathbb{E}\big(f(\tau'')|\tau'\big), g\big)\,;$$

thus,

$$\mathbb{E}\big(f(\tau'')|\tau'\big) = (\rho^{\mathbf{N}} f)(\tau')\,. \tag{1.3}$$

(Our $\rho^{\mathbf{N}}$ is $T_\eta = Q_\varepsilon$ of [3] with $\eta = \rho$, $\varepsilon = (1-\rho)/2$.) (In fact, let $\mathbf{N} f = \sum_n n \tilde f_n$, then $-\mathbf{N}$ is the generator of a Markov process on $\{-1,+1\}^n$; $\exp(-t\mathbf{N})$ is its semigroup; note that $\rho^{\mathbf{N}}$ is of the form $\exp(-t\mathbf{N})$. The Markov process is quite simple: during dt, each coordinate flips with the probability $\frac12 dt + o(dt)$. However, we do not need it.) Note also that $\mathbb{E}\big(|f_n(\tau'') - (\rho^{\mathbf{N}} f_n)(\tau')|^2 \,\big|\, \tau'\big)$ is the conditional variance $\operatorname{Var}\big(f_n(\tau'')\big|\tau'\big)$, and its mean value (over all τ') is

$$\mathbb{E}\operatorname{Var}\big(f_n(\tau'')\big|\tau'\big) = \|f_n\|^2 - \|\rho^{\mathbf{N}} f_n\|^2 = \big((\mathbf{1} - \rho^{2\mathbf{N}}) f_n, f_n\big)\,. \tag{1.4}$$

Note also that the operator $0^{\mathbf{N}} = \lim_{\rho\to 0} \rho^{\mathbf{N}}$ is the projection onto the one-dimensional space of constants, $f \mapsto (\mathbb{E} f) \cdot \mathbf{1}$.
Stability of $F = \big(f_n\big)_{n=1}^\infty$ is equivalent to:

- $\|\rho^{\mathbf{N}} f_n - f_n\| \xrightarrow[\rho\to 1]{} 0$ uniformly in n;
- $(\rho^{\mathbf{N}} f_n, f_n) \xrightarrow[\rho\to 1]{} \|f\|^2$ uniformly in n;

- $\|f_n\|^2 - \|\rho^{\mathbf{N}} f_n\|^2 \xrightarrow[\rho\to 1]{} 0$ uniformly in n.

Sensitivity of F is equivalent to:

- $\|(\rho^{\mathbf{N}} - 0^{\mathbf{N}})f_n\| \xrightarrow[n\to\infty]{} 0$ for some (or every) $\rho \in (0,1)$;
- $\big((\rho^{\mathbf{N}} - 0^{\mathbf{N}})f_n, f_n\big) \xrightarrow[n\to\infty]{} 0$ for some (or every) $\rho \in (0,1)$.

Combining these facts with the probabilistic interpretation (1.2), (1.3), (1.4) of $\rho^{\mathbf{N}}$ we see that

- F is stable iff $\mathbb{E}\big(\overline{f_n(\tau')}f_n(\tau'')\big) \xrightarrow[\rho\to 1]{} \mathbb{E}|f_n(\tau)|^2$ uniformly in n
  or, equivalently, $\mathbb{E}\big(\operatorname{Var}(f_n(\tau'')|\tau')\big) \xrightarrow[\rho\to 1]{} 0$ uniformly in n;
- F is sensitive iff $\mathbb{E}\big(\overline{f_n(\tau')}f_n(\tau'')\big) - \big|\mathbb{E}f_n(\tau)\big|^2 \xrightarrow[n\to\infty]{} 0$ for some (or every) $\rho \in (0,1)$ or, equivalently, $\mathbb{E}\,\big|\mathbb{E}(f(\tau'')|\tau') - \mathbb{E}f\big|^2 \xrightarrow[n\to\infty]{} 0$ for some (or every) $\rho \in (0,1)$.

These are versions of definitions introduced in [3, Sect. 1.1, 1.4].

2 Stability and sensitivity on trees

A branch of the n-level binary tree can be written as a sequence of sequences $()$, (τ_1), (τ_1,τ_2), (τ_1,τ_2,τ_3), ..., $(\tau_1,\dots,\tau_n)$. Branches correspond to leaves $(\tau_1,\dots,\tau_n) \in \{-1,+1\}^n$. Automorphisms of the tree can be described as maps $A : \{-1,+1\}^n \to \{-1,+1\}^n$ of the form

$$A(\tau_1,\dots,\tau_n) = \big(a()\tau_1, a(\tau_1)\tau_2, a(\tau_1,\tau_2)\tau_3, \dots, a(\tau_1,\dots,\tau_{n-1})\tau_n\big)$$

for arbitrary functions $a : \cup_{m=1}^n \{-1,+1\}^{m-1} \to \{-1,+1\}$. (Thus, the tree has $2^1 \cdot 2^2 \cdot 2^4 \cdot \ldots \cdot 2^{2^{n-1}} = 2^{2^n-1}$ automorphisms, while the cube $\{-1,+1\}^n$ has only $2^n n!$ automorphisms.) Here is an example of a tree automorphism (far from being a cube automorphism):

$$(\tau_1,\dots,\tau_n) \mapsto \big(\tau_1, \tau_1\tau_2, \dots, \tau_1\dots\tau_n\big)\,.$$

The function $f_n(\tau_1,\dots,\tau_n) = \frac{1}{\sqrt{n}}(\tau_1 + \dots + \tau_n)$ satisfies $S_1^1(f_n) = 1$, $S_2^\infty(f_n) = 0$. However, the function $g_n(\tau_1,\dots,\tau_n) = \frac{1}{\sqrt{n}}\big(\tau_1 + \tau_1\tau_2 + \dots + \tau_1\dots\tau_n\big)$ satisfies $S_1^m(g_n) = \min\big(\frac{m}{n}, 1\big)$, $S_m^\infty(g_n) = \max\big(\frac{n-m+1}{n}, 0\big)$. According to the definitions of Sect. 1, $(f_n)_{n=1}^\infty$ is stable, but $(g_n)_{n=1}^\infty$ is sensitive. We see that the definitions are not tree-invariant. A straightforward way to tree-invariance is used in the following definition of "tree stability" and "tree sensitivity". From now on, stability and sensitivity of Sect. 1 will be called "cube stability" and "cube sensitivity".

2.1 Definition (a) A sequence $(f_n)_{n=1}^\infty$ of functions $f_n : \{-1,+1\}^n \to \mathbb{C}$ is *tree stable,* if there exists a sequence of tree automorphisms $A_n : \{-1,+1\}^n \to \{-1,+1\}^n$ such that the sequence $\big(f_n \circ A_n\big)_{n=1}^\infty$ is cube stable.
(b) The sequence $(f_n)_{n=1}^\infty$ is *tree sensitive,* if $\big(f_n \circ A_n\big)_{n=1}^\infty$ is cube sensitive for every sequence (A_n) of tree automorphisms.

The definition can be formulated in terms of $f_n(A_n(\tau'))$ and $f_n(A_n(\tau''))$ where (τ',τ'') is a ρ-correlated pair of random points of the cube $\{-1,+1\}^n$. Equivalently, we may consider $f_n(\tau')$ and $f_n(\tau'')$ where τ',τ'' are such that for some A_n, $(A_n\tau', A_n\tau'')$ is a ρ-correlated pair. That is,

$$\mathbb{E}\big(\tau'_m \big| \tau'_1,\tau''_1,\dots,\tau'_{m-1},\tau''_{m-1}\big) = \mathbb{E}\big(\tau''_m \big| \tau'_1,\tau''_1,\dots,\tau'_{m-1},\tau''_{m-1}\big) = 0\,, \tag{2.2}$$

$$\mathbb{E}\big(\tau'_m\tau''_m \big| \tau'_1,\tau''_1,\dots,\tau'_{m-1},\tau''_{m-1}\big) = a(\tau'_1,\dots,\tau'_{m-1})a(\tau''_1,\dots,\tau''_{m-1})\rho\,, \tag{2.3}$$

where $a : \cup_{m=1}^n \{-1,+1\}^{m-1} \to \{-1,+1\}$. On the other hand, consider an arbitrary $\{-1,+1\}^n \times \{-1,+1\}^n$-valued random variable (τ',τ'') satisfying (2.2) (which implies that each one of τ',τ'' is uniform on $\{-1,+1\}^n$), but maybe not (2.3), and define

$$\rho_{\max}(\tau',\tau'') = \max_{m=1,\dots,n} \max \big| \mathbb{E}\big(\tau'_m\tau''_m \big| \tau'_1,\tau''_1,\dots,\tau'_{m-1},\tau''_{m-1}\big) \big|\,, \tag{2.4}$$

where the internal maximum is taken over all possible values of $(\tau'_1,\tau''_1,\dots,\tau'_{m-1},\tau''_{m-1})$. The joint distribution of τ' and τ'' is a probability measure μ on $\{-1,+1\}^n \times \{-1,+1\}^n$, and we denote $\rho_{\max}(\tau',\tau'')$ by $\rho_{\max}(\mu)$. Given $f,g : \{-1,+1\}^n \to \mathbb{C}$, we denote $\mathbb{E}\overline{f(\tau')}g(\tau'')$ by $\langle f|\mu|g\rangle$.

2.5 Definition A sequence $(f_n)_{n=1}^\infty$ of functions $f_n : \{-1,+1\}^n \to \mathbb{C}$, satisfying $0 < \liminf_{n\to\infty} \|f_n\| \le \limsup_{n\to\infty} \|f_n\| < \infty$, is *cosy,* if for any $\varepsilon > 0$ there is a sequence $(\mu_n)_{n=1}^\infty$, μ_n being a probability measure on $\{-1,+1\}^n \times \{-1,+1\}^n$, such that $\limsup_{n\to\infty} \rho_{\max}(\mu_n) < 1$ and $\limsup_{n\to\infty} \big(\|f_n\|^2 - \langle f_n|\mu_n|f_n\rangle\big) < \varepsilon$.

2.6 Lemma Every tree stable sequence is cosy.

PROOF. Let (f_n) be tree stable. Take tree automorphisms A_n such that $(f_n \circ A_n)$ is cube stable. We have $\mathbb{E}\big(\overline{f_n(A_n(\tau'))} f_n(A_n(\tau''))\big) \xrightarrow[\rho\to1]{} \mathbb{E}|f_n(\tau)|^2$ uniformly in n. Here τ',τ'' are ρ-correlated. The joint distribution $\mu_n(\rho)$ of $A_n(\tau')$ and $A_n(\tau'')$ satisfies $\rho_{\max}(\mu_n(\rho)) \le \rho$ due to (2.3). Also, $\langle f_n|\mu_n|f_n\rangle \xrightarrow[\rho\to1]{} \|f_n\|^2$ uniformly in n, which means that $\sup_n\big(\|f_n\|^2 - \langle f_n|\mu_n|f_n\rangle\big) \to 0$ for $\rho \to 1$. □

Is there a cosy but not tree stable sequence? We do not know. The conditional correlation given by (2.3) is not only $\pm\rho$, it is also factorizable (a function of τ' times the same function of τ''), which seems to be much stronger than just $\rho_{\max}(\mu) \le \rho$.

3 Hypercontractivity

Let (τ',τ'') be a ρ-correlated pair of random points of the cube $\{-1,+1\}^n$. Then for every $f,g : \{-1,+1\}^n \to \mathbb{R}$

$$\big|\mathbb{E}f(\tau')g(\tau'')\big|^{1+\rho} \le \big(\mathbb{E}|f(\tau')|^{1+\rho}\big)\big(\mathbb{E}|g(\tau'')|^{1+\rho}\big)\,, \tag{3.1}$$

which is a discrete version of the celebrated hypercontractivity theorem pioneered by Nelson (see [7, Sect. 3]). For a proof, see [1]; there, following Gross [6], the inequality is proved for $n = 1$ (just two points, $\{-1,+1\}$) [1, Prop. 1.5], which is enough due to tensorization [1, Lemma 1.3]. (See also [3, Lemma 2.4].) The case of f, g taking on two values 0 and 1 only is especially important:

$$\mathbb{P}^{1+\rho}\big(\tau' \in S' \,\&\, \tau'' \in S''\big) \le \mathbb{P}(\tau' \in S')\mathbb{P}(\tau'' \in S'') = \frac{|S'|}{2^n} \cdot \frac{|S''|}{2^n}$$

for any $S', S'' \subset \{-1,+1\}^n$. Note that $\rho = 0$ means independence,[1] while $\rho = 1$ is trivial: $\mathbb{P}^2(\dots) \le \big(\min(\mathbb{P}(S'), \mathbb{P}(S''))\big)^2 \le \mathbb{P}(S')\mathbb{P}(S'')$.

For a probability measure μ on $\{-1,+1\}^n \times \{-1,+1\}^n$ we denote by $\langle g|\mu|f\rangle$ the value $\mathbb{E}\big(f(\tau')g(\tau'')\big)$, where $(\tau', \tau'') \sim \mu$. The hypercontractivity (3.1) may be written as $\big|\langle g|\mu|f\rangle\big| \le \|f\|_{1+\rho}\|g\|_{1+\rho}$, where $\mu = \mu(\rho)$ is the distribution of a ρ-correlated pair. The class of μ that satisfy the inequality (for all f, g) is invariant under transformations of the form $A \times B$, where $A, B : \{-1,+1\}^n \to \{-1,+1\}^n$ are arbitrary invertible maps (since such maps preserve $\|\cdot\|_{1+\rho}$). In particular, all measures of the form (2.2–2.3) fit.

Can we generalize the statement for all μ such that $\rho_{\max}(\mu) \le \rho$? The approach of Gross, based on tensorization, works on cubes (and other products), not trees. Fortunately, we have another approach, found by Neveu [8], that works also on trees.

3.2 Lemma For every $r \in [\frac12, 1]$, $x, y \in [0,1]$, and $\rho \in [-\frac{1-r}{r}, \frac{1-r}{r}]$,

$$\begin{aligned}(1+\rho)(1-x)^r(1-y)^r + (1-\rho)(1-x)^r(1+y)^r +{}& \\ + (1-\rho)(1+x)^r(1-y)^r + (1+\rho)(1+x)^r(1+y)^r &\le 4\,.\end{aligned}$$

Proof. The left hand side is linear in ρ with the coefficient $\big((1+x)^r - (1-x)^r\big)\big((1+y)^r - (1-y)^r\big) \ge 0$. Therefore, it suffices to prove the inequality for $\rho = \frac{1-r}{r}$, $r \in (\frac12, 1)$ (the cases $r = \frac12$ and $r = 1$ follow by continuity). Assume the contrary, then the continuous function f_r on $[0,1] \times [0,1]$, defined by

$$\begin{aligned} f_r(x,y) = \frac1r (1-x)^r(1-y)^r + \frac{2r-1}{r}(1-x)^r(1+y)^r &+ \\ + \frac{2r-1}{r}(1+x)^r(1-y)^r + \frac1r (1+x)^r(1+y)^r &\,, \end{aligned}$$

has a global maximum $f_r(x_0, y_0) > 4$ for some $r \in (\frac12, 1)$. The case $x_0 = y_0 = 0$ is excluded (since $f_r(0,0) = 4$). Also, $x_0 \ne 1$ (since $\frac{\partial}{\partial x}\big|_{x=1-} f_r(x,y) = -\infty$) and $y_0 \ne 1$. The new variables

$$u = \frac{1+x}{1-x} \in [1, \infty)\,, \quad v = \frac{1+y}{1-y} \in [1, \infty)$$

will be useful. We have

$$\frac{1+x}{(1-x)^r(1-y)^r}\frac{\partial}{\partial x} f_r(x,y) = u^r v^r - u - (2r-1)(uv^r - u^r)\,. \tag{3.3}$$

For $u = 1$, $v > 1$ the right hand side is $2(1-r)(v^r - 1) > 0$; therefore $x_0 \ne 0$ (since $(x_0, y_0) \ne (0,0)$), and similarly $y_0 \ne 0$. So, (x_0, y_0) is an interior point of $[0,1] \times [0,1]$. The corresponding $u_0, v_0 \in (1, \infty)$ satisfy $u_0^r v_0^r - u_0 - (2r-1)(u_0 v_0^r - u_0^r) = 0$. By subtracting the same expression with v_0 switched with u_0, which also vanishes, we get

$$v_0 - u_0 + (2r-1)(u_0^r v_0 - u_0 v_0^r + u_0^r - v_0^r) = 0\,.$$

[1] Equality results from the inequality applied to complementary sets.

Aiming to conclude that $u_0 = v_0$, consider the function $u \mapsto v_0 - u + (2r-1)(u^r v_0 - u v_0^r + u^r - v_0^r)$ on $[1,\infty)$. It is concave, and positive when $u = 1$, since $v_0 - 1 + (2r-1)(v_0 - 2v_0^r + 1) \ge v_0 - 1 + (2r-1)(v_0 - 2v_0 + 1) = (v_0 - 1)(2 - 2r)$. Therefore, the function cannot vanish more than once, and $u = v_0$ is its unique root. So, $u_0 = v_0$.
It follows from (3.3) that

$$\frac{1+x}{(1-x)^{2r}} \cdot \frac{1}{2}\frac{\partial}{\partial x} f(x,x) = u^{2r} - u - (2r-1)(u^{r+1} - u^r)\,,$$

therefore u_0 is a root of the equation $u^{2r-1} - 1 - (2r-1)(u^r - u^{r-1}) = 0$, different from the evident root $u = 1$. However, the function $u \mapsto u^{2r-1} - 1 - (2r-1)(u^r - u^{r-1})$ is strictly monotone, since

$$\frac{1}{2r-1}\frac{\partial}{\partial u}(\dots) = u^{2r-2} - ru^{r-1} + (r-1)u^{r-2} = u^{r-2}(u^r - ru + r - 1) < 0$$

due to the inequality $u^r \le 1 + r(u-1)$ (which follows from concavity of u^r). The contradiction completes the proof. $\square$

3.4 Theorem Let $\rho \in [0,1]$, and μ be a probability measure on $\{-1,+1\}^n \times \{-1,+1\}^n$ such that[2] $\rho_{\max}(\mu) \le \rho$. Then for every $f, g : \{-1,+1\}^n \to \mathbb{C}$

$$\big|\langle g|\mu|f\rangle\big| \le \|f\|_{1+\rho}\|g\|_{1+\rho}\,.$$

PROOF. Consider random points τ', τ'' of $\{-1,+1\}^n$ such that $(\tau', \tau'') \sim \mu$. We have two (correlated) random processes $\tau'_1, \dots, \tau'_n$ and $\tau''_1, \dots, \tau''_n$. Consider the random variables

$$M'_n = |f(\tau'_1, \dots, \tau'_n)|^{1/r}\,, \quad M''_n = |g(\tau''_1, \dots, \tau''_n)|^{1/r}\,,$$

and the corresponding martingales

$$\begin{aligned} M'_m &= \mathbb{E}\big(M'_n \big| \tau'_1, \tau''_1, \dots, \tau'_m, \tau''_m\big) = \mathbb{E}\big(M'_n \big| \tau'_1, \dots, \tau'_m\big)\,, \\ M''_m &= \mathbb{E}\big(M''_n \big| \tau'_1, \tau''_1, \dots, \tau'_m, \tau''_m\big) = \mathbb{E}\big(M''_n \big| \tau''_1, \dots, \tau''_m\big) \end{aligned}$$

for $m = 0, 1, \dots, n$; the equalities for conditional expectations follow from (2.2). For any $m = 1, \dots, n$ and any values of $\tau'_1, \tau''_1, \dots, \tau'_{m-1}, \tau''_{m-1}$ consider the conditional distribution of the pair (M'_m, M''_m). It is concentrated at four points that can be written as[3] $\big((1\pm x)M'_{m-1}, (1\pm y)M''_{m-1}\big)$. The first "$\pm$" depends only on τ'_m, the second on τ''_m (given the past); each of them is "$-$" or "$+$" equiprobably. They have some correlation coefficient lying between $(-\rho)$ and ρ. Lemma 3.2 gives

$$4\mathbb{E}\bigg(\Big(\frac{M'_m}{M'_{m-1}}\frac{M''_m}{M''_{m-1}}\Big)^r \bigg| \dots\bigg) \le 4\,,$$

where $r = \frac{1}{1+\rho}$. Thus, $\mathbb{E}\big((M'_m M''_m)^r \big| \dots\big) \le (M'_{m-1}M''_{m-1})^r$, which means that the process $(M'_m M''_m)^r$ is a supermartingale. Therefore, $\mathbb{E}(M'_n M''_n)^r \le (M'_0 M''_0)^r$, that is,

$$\begin{aligned} \mathbb{E}|f(\tau'_1, \dots, \tau'_n) g(\tau''_1, \dots, \tau''_n)| &\le \big(\mathbb{E}|f(\tau'_1, \dots, \tau'_n)|^{1/r}\big)^r \cdot \big(\mathbb{E}|g(\tau''_1, \dots, \tau''_n)|^{1/r}\big)^r \\ &= \|f\|_{1+\rho}\|g\|_{1+\rho}\,. \end{aligned}$$

$\square$

[2] It is assumed that μ satisfies (2.2); $\rho_{\max}$ was defined only for such measures.
[3] Of course, x and y depend on $\tau'_1, \tau''_1, \dots, \tau'_{m-1}, \tau''_{m-1}$.

4 The main result

Return to the spider walk (Fig. 1b). It may be treated as a complex-valued martingale Z (Fig. 2a), starting at the origin. Take each step to have length 1. The set Ω_n^{spider} of all n-step trajectories of Z can be identified with the set of leaves of a binary tree. The endpoint $Z_n = Z_n(\omega)$ of a trajectory $\omega \in \Omega_n^{\text{spider}}$ is a complex-valued function on Ω_n^{spider}. Taking into account that $\mathbb{E}|Z_n|^2 = n$, we ask about tree stability of the sequence $\big(Z_n/\sqrt{n}\big)_{n=1}^\infty$.

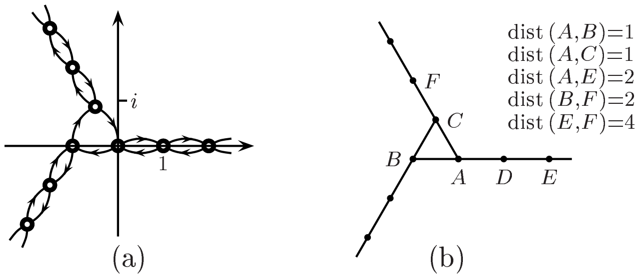


Figure 2: (a) the spider walk as a complex-valued martingale; (b) combinatorial distance.

4.1 Theorem The sequence $\big(Z_n/\sqrt{n}\big)_{n=1}^\infty$ is non-cosy.

By Lemma 2.6 it follows that the sequence $\big(Z_n/\sqrt{n}\big)_{n=1}^\infty$ is not tree stable. Recently, M. Emery and J. Warren found that some tree sensitive sequences result naturally from their constructions.
In contrast to the spider walk, the simple walk (Fig. 1a) produces a sequence $\big((\tau_1 + \dots + \tau_n)/\sqrt{n}\big)_{n=1}^\infty$ that evidently is cube stable, therefore tree stable, therefore cosy.

4.2 Lemma (a) $\limsup_{n\to\infty} \sqrt{n}\mathbb{P}(Z_n = 0) < \infty$.
(b) $\liminf_{n\to\infty} \big(n^{-1/2} \sum_{k=1}^n \mathbb{P}(Z_k = 0)\big) > 0$.

The proof is left to the reader. Both (a) and (b) hold for each node of our graph, not just 0. In fact, the limit exists,

$$\lim_{n\to\infty} \big(n^{1/2}\mathbb{P}(Z_n = 0)\big) = \frac{1}{2} \lim_{n\to\infty} \big(n^{-1/2} \sum_{k=1}^n \mathbb{P}(Z_k = 0)\big) \in (0,\infty)\,,$$

but we do not need it.

Proof of the theorem. Let μ_n be a probability measure on $\Omega_n^{\text{spider}} \times \Omega_n^{\text{spider}}$ such that[4] $\rho_{\max}(\mu) \le \rho$, $\rho \in (0,1)$; we'll estimate $\langle Z_n|\mu_n|Z_n\rangle$ from above in terms of ρ. We have two (correlated) copies $\big(Z'_k\big)_{k=1}^n$, $\big(Z''_k\big)_{k=1}^n$ of the martingale $\big(Z_k\big)_{k=1}^n$. Consider the combinatorial distance (see Fig. 2b)

$$D_k = \operatorname{dist}(Z'_k, Z''_k)\,.$$

Conditionally, given the past $(Z'_1, Z''_1, \dots, Z'_{m-1}, Z''_{m-1})$, we have two *equiprobable* values for Z'_m, and two *equiprobable* values for Z''_m; the two binary choices are correlated, their correlation

[4]It is assumed that μ satisfies (2.2); $\rho_{\max}$ was defined only for such measures.

lying in $[-\rho, \rho]$. The four possible values for (Z'_m, Z''_m) lead usually to three possible values $D_{m-1} - 2$, D_{m-1}, $D_{m-1} + 2$ for D_m, see Fig. 3a; their probabilities depend on the correlation, but the (conditional) expectation of D_m is equal to D_{m-1} irrespective of the correlation. Sometimes, however, a different situation appears, see Fig. 3b; here the conditional expectation of D_m is equal to $D_{m-1} + 1/2$ rather than D_{m-1}. That happens when Z''_{m-1} is situated at the beginning of a ray (any one of our three rays) and Z'_{m-1} is on the same ray, outside the central triangle Δ (ABC on Fig. 2b). In that case[5] we set $L_{m-1} = 1$, otherwise $L_{m-1} = 0$. We do not care about the case when Z'_{m-1}, Z''_{m-1} are both on Δ; this case may be neglected due to hypercontractivity, as we'll see soon. Also, the situation where $Z'_{m-1} = Z''_{m-1}$ may occur, and then $\mathbb{E}\big(D_m \big| D_{m-1}\big) \ge D_{m-1}$.

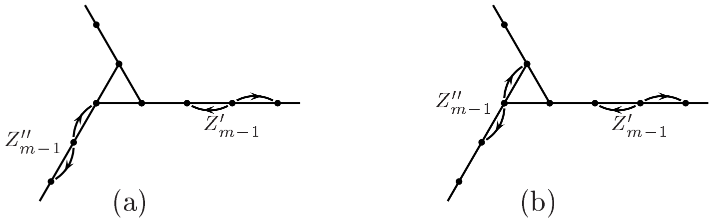


Figure 3: (a) the usual case, $L = 0$: in the mean, D remains the same; (b) the case of $L = 1$: in the mean, D increase by $1/2$. More cases exist, but D never decreases in the mean.

Theorem 3.4, applied to appropriate indicators, gives $\mathbb{P}^{1+\rho}\big(Z'_k \in \Delta \,\&\, Z''_k \in \Delta\big) \le \mathbb{P}\big(Z'_k \in \Delta\big) \cdot \mathbb{P}\big(Z''_k \in \Delta\big)$, that is,

$$\mathbb{P}\big(Z'_k \in \Delta \,\&\, Z''_k \in \Delta\big) \le \big(\mathbb{P}(Z_k \in \Delta)\big)^{\frac{2}{1+\rho}}$$

for all $k = 0, \dots, n$. Combining it with Lemma 4.2 (a) we get

$$\sum_{k=0}^{n} \mathbb{P}\big(Z'_k \in \Delta \,\&\, Z''_k \in \Delta\big) \le \varepsilon_n(\rho) \cdot \sqrt{n} \tag{4.3}$$

for some $\varepsilon_n(\rho)$ such that $\varepsilon_n(\rho) \xrightarrow[n\to\infty]{} 0$ for every $\rho \in (0,1)$, and $\varepsilon_n(\rho)$ does not depend on μ as long as $\rho_{\max}(\mu) \le \rho$.
Now we are in position to show that

$$\sum_{k=0}^{n} \mathbb{P}\big(L_k = 1\big) \ge c_0 \sqrt{n} \tag{4.4}$$

for $n \ge n_0(\rho)$; here $n_0(\rho)$ and $c_0 > 0$ do not depend on μ. First, Lemma 4.2 (b) shows that $\mathbb{P}(Z''_k = 0)$ is large enough. Second, (4.3) shows that $\mathbb{P}(Z''_k = 0 \,\&\, Z'_k \notin \Delta)$ is still large enough. The same holds for $\mathbb{P}(Z''_k = 0 \,\&\, Z'_k \notin \Delta_{+2})$, where Δ_{+2} is the (combinatorial) 2-neighborhood of Δ. Last, given that $Z''_k = 0$ and $Z'_k \notin \Delta_{+2}$, we have a not-so-small (in fact, $\ge 1/4$) conditional probability that $L_k + L_{k+1} + L_{k+2} > 0$. This proves (4.4).

[5] There is a symmetric case (Z'_{m-1} at the beginning ...), but we do not use it.

The process $\big(D_m - \frac12 \sum_{k=0}^{m-1} L_k\big)_{m=0}^n$ is a submartingale (that is, increases in the mean). Therefore, using (4.4),

$$\mathbb{E}D_n \ge \frac12 \sum_{k=0}^{n-1} \mathbb{P}(L_k = 1) \ge \frac12 c_0 \sqrt{n}$$

for $n \ge n_0(\rho)$. Note that $D_n = \operatorname{dist}(Z'_n, Z''_n) \le C_1 |Z'_n - Z''_n|$ for some absolute constant C_1. We have

$$\big(\mathbb{E}|Z'_n - Z''_n|^2\big)^{1/2} \ge \mathbb{E}|Z'_n - Z''_n| \ge C_1^{-1}\mathbb{E}D_n \ge \frac12 C_1^{-1} c_0 \sqrt{n}$$

and

$$\|Z_n\|^2 - \langle Z_n|\mu_n|Z_n\rangle = \frac12 \mathbb{E}|Z'_n - Z''_n|^2 \ge \frac14 C_1^{-2} c_0^2 n$$

for $n \ge n_0(\rho)$; so,

$$\limsup_{n\to\infty} \left(\left\| \frac{Z_n}{\sqrt{n}} \right\|^2 - \left\langle \frac{Z_n}{\sqrt{n}} \middle| \mu_n \middle| \frac{Z_n}{\sqrt{n}} \right\rangle \right) \ge \frac{c_0^2}{4C_1^2}$$

irrespective of ρ, which means non-cosiness. □

5 Connections to continuous models

Theorem 4.1 (non-cosiness) is a discrete counterpart of [9, Th. 4.13]. A continuous complex-valued martingale $Z(t)$ considered there, so-called Walsh's Brownian motion, is the limit of our $\big(Z_{nt}/\sqrt{n}\big)$ when $n \to \infty$. The constants c_0 and C_1 used in the proof of Theorem 4.1 can be improved (in fact, made optimal) by using explicit calculations for Walsh's Brownian motion. Cosiness for the simple walk is a discrete counterpart of [9, Lemma 2.5].

Theorem 3.3 (hypercontractivity on trees) is a discrete counterpart of [9, Lemma 6.5]. However, our use of hypercontractivity when proving non-cosiness follows [2, pp. 278–280]. It is possible to estimate $\mathbb{P}(Z'_k \in \Delta \,\&\, Z''_k \in \Delta)$ without hypercontractivity, following [5] or [9, Sect. 4].

Cosiness, defined in Def. 2.5, is a discrete counterpart of the notion of cosiness introduced in [9, Def. 2.4]. Different variants of cosiness (called I-cosiness and D-cosiness) are investigated by Émery, Schachermayer, and Beghdadi-Sakrani, see [4] and references therein. See also Warren [13].

Noise stability and noise sensitivity, introduced in [3], have their continuous counterparts, see [10, 11]. Stability corresponds to white noises, sensitivity to black noises. Mixed cases (neither stable nor sensitive, see [3, end of Sect. 1.4]) correspond to noises that are neither white nor black (as in [14]).

References

[1] D. Bakry, "L'hypercontractivité et son utilisation en théorie des semigroups", *Lect. Notes Math* (Lectures on probability theory), Springer, Berlin, **1581** (1994), 1–114.

[2] M.T. Barlow, M. Émery, F.B. Knight, S. Song, M. Yor, "Autour d'un théorème de Tsirelson sur des filtrations browniennes et non browniennes", *Lect. Notes Math* (Séminaire de Probabilités XXXII), Springer, Berlin, **1686** (1998), 264–305.

[3] I. Benjamini, G. Kalai, O. Schramm, "Noise sensitivity of Boolean functions and applications to percolation", preprint math.PR/9811157.

[4] M. Émery, "Remarks on an example studied by A. Vershik and M. Smorodinsky", Manuscript, 1998.

[5] M. Émery, M. Yor, "Sur un théorème de Tsirelson relatif à des mouvements browniens corrélés et à la nullité de certains temps locaux", *Lect. Notes Math.* (Séminaire de Probabilités XXXII), Springer, Berlin, **1686** (1998), 306–312.

[6] L. Gross, "Logarithmic Sobolev inequalities", *Amer. J. Math.* **97**:4 (1976), 1061–1083.

[7] E. Nelson, "The free Markoff field", *J. Funct. Anal.* **12** (1973), 211–227.

[8] J. Neveu, "Sur l'espérance conditionelle par rapport à un mouvement brownien", *Ann. Inst. H. Poincaré*, Sect. B, **12**:2 (1976), 105–109.

[9] B. Tsirelson, "Triple points: from non-Brownian filtrations to harmonic measures," *Geom. Funct. Anal. (GAFA)* **7** (1997), 1096–1142.

[10] B. Tsirelson, "Scaling limit of Fourier-Walsh coefficients (a framework)", preprint math.PR/9903121.

[11] B. Tsirelson, "Noise sensitivity on continuous products: an answer to an old question of J. Feldman", preprint math.PR/9907011.

[12] B.S. Tsirelson, A.M. Vershik, "Examples of nonlinear continuous tensor products of measure spaces and non-Fock factorizations", *Reviews in Mathematical Physics* **10**:1 (1998), 81–145.

[13] J. Warren, "On the joining of sticky Brownian motion", *Lect. Notes Math.* (Séminaire de Probabilités XXXIII), Springer, Berlin (to appear).

[14] J. Warren, "The noise made by a Poisson snake", Manuscript, Univ. de Pierre et Marie Curie, Paris, Nov. 1998.