Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 2, 427–442
DOI: 10.1214/14-BJPS256

A note on the Bramson–Kalikow process

Sacha Friedli

Universidade Federal de Minas Gerais

Abstract. We consider discrete-time stationary processes with long-range dependencies, $X_n \in \{\pm 1\}$, $n \in \mathbb{Z}$, specified by a regular attractive g-function, similar to those considered by Bramson and Kalikow [*Israel J. Math.* **84** (1993) 153–160]. We give an explicit set of conditions that imply the existence of at least two distinct processes specified by the same g-function, and consider a few examples that emphasize the role played by the smoothness of the majority rule at the origin.

1 Introduction

Consider a measurable map $g : \{\pm 1\}^{\mathbb{N}} \to [0, 1]$, called g-function. A stationary processes $X = (X_n)_{n \in \mathbb{Z}}$ with $X_n \in \{\pm 1\}$, is said to be specified by g if

$$P(X_{n+1} = +1 | X_n = \sigma_n, X_{n-1} = \sigma_{n-1}, \dots) = g(\sigma_n, \sigma_{n-1}, \dots) \qquad P\text{-a.s.} \tag{1}$$

Let $\mathrm{var}_k(g)$ denote the variation of g of order k, that is,

$$\mathrm{var}_k(g) := \sup\{|g(\sigma) - g(\sigma')| : \sigma_i = \sigma'_i, 1 \le i \le k\}.$$

It is well known that if g is regular, that is if $\mathrm{var}_k(g) \to 0$ when $k \to \infty$, and if there exists $\varepsilon_* \in (0, \frac{1}{2})$ such that $g(\sigma) \in [\varepsilon_*, 1 - \varepsilon_*]$ for all σ, then there always exists at least one process specified by g.

Most of the existing literature on this kind of process is concerned with finding additional conditions on g which imply uniqueness of X. Uniqueness is not our main concern here, but we mention two important contributions. In their pioneering paper (Doeblin and Fortet, 1937), Doeblin and Fortet showed that uniqueness holds when $\mathrm{var}_k(g) \in \ell^1$. More recently, Johansson and Öberg (2003) obtained the same conclusion under the weaker assumption that $\mathrm{var}_k(g) \in \ell^2$.

As in the theory of phase transitions in equilibrium statistical mechanics, it is natural to ask whether non-uniqueness is possible: can two distinct stationary processes be specified by the same regular g-function? The first result on non-uniqueness was obtained by Bramson and Kalikow (1993), who considered g-functions of the following kind. Let $1 \le m_1 < m_2 < \cdots$ be an increasing sequence of odd integers, and set

$$g(\sigma_1, \sigma_2, \dots) := \sum_{k \ge 1} p_k \varphi\left(\frac{1}{m_k} \sum_{j=1}^{m_k} \sigma_j\right), \tag{2}$$

Key words and phrases. Long-range, g-function, complete connections, phase transition, majority rule.

Received March 2014; accepted August 2014.

where $p_k \geq 0$, $\sum_{k\geq 1} p_k = 1$, and $\varphi : [-1, 1] \to [-1, 1]$ is non-decreasing. The function used in Bramson and Kalikow (1993) was the pure majority rule, defined by

$$\varphi_{\mathrm{PMR}}(s) := \begin{cases} 1 - \varepsilon_*, & \text{if } s > 0, \\ \varepsilon_*, & \text{if } s < 0. \end{cases} \tag{3}$$

With this choice, (1) is interpreted by saying that to determine the state of X_{n+1} for a given past, one looks at all variables X_j lying at a distance $\leq m_k$ in the past of $n + 1$, with probability p_k, and with probability $1 - \varepsilon_*$, X_{n+1} is assigned the value taken by the sign of the majority of the variables in the block $[n - m_k + 1, n]$ (since m_k is odd, the majority is always well defined).

Bramson and Kalikow showed that when using φ_{PMR}, it becomes possible to choose the sequences (p_k) and (m_k) such that at least two distinct stationary processes satisfy (1). The technique used in the proof consists in first fixing (p_k), and then constructing (m_k) inductively: at the kth step, the pointwise ergodic theorem for Markov chains is used to guarantee the existence of m_{k+1} as a function of $m_1, \ldots, m_k$ and $p_1, \ldots, p_k$. An aftermath is that the sequence (m_k) is lacunary, $m_{k+1} \gg m_k$, and diverges in an uncontrolled way, but the g-function obtained is known to exhibit non-uniqueness. An important ingredient of the proof is that the corresponding g-function is attractive: $g(\sigma) \leq g(\sigma')$ when $\sigma \leq \sigma'$,[1] which implies that the presence of $+1$s in the past favorizes their presence in the near future too.

More recently, Berger, Hoffman and Sidoravicius (2003) obtained non-uniqueness for a different class of (non-attractive) g-functions, whose analysis was pushed further in Dias and Friedli (2015) (see the remarks below and at the end of the paper). The proof in Berger et al. (2003) also showed that the criterion of Johansson and Öberg is optimal, in the following sense: for all $\varepsilon > 0$ there exists a g-function with $\mathrm{var}_k(g) \in \ell^{2+\varepsilon}$, for which non-uniqueness holds.

In this note, we give a closer look at the inductive construction of Bramson–Kalikow for attractive g-functions, and give a set of conditions (see Theorem 5.1) leading to an algorithm for choosing m_{k+1} as a function of m_k. Our treatment of the induction step uses a concentration inequality (Samson, 2000) rather than the ergodic theorem, together with a convergence result from Doeblin and Fortet (1937). Our method provides a concentration inequality for processes specified by g-functions with summable variations (Theorem 4.2), which is of independent interest. Although it does not yet allow to go beyond the realm of lacunary sequences, it still provides a more accurate criterium that can be tested to determine whether a given g-functions exhibits non-uniqueness.

[1] The partial order on $\{\pm 1\}^{\mathbb{N}}$ is the usual one: $\sigma \leq \sigma'$ if and only if $\sigma_i \leq \sigma'_i$ for all $i \in \mathbb{N}$.

But we will also be interested in the role played by the behavior of φ at $s = 0$. As a matter of fact, the Bramson–Kalikow mechanism, as presented in Bramson and Kalikow (1993), is possible due to the discontinuity of φ_{PMR} at $s = 0$. This discontinuity implies that small local fluctuations can have important consequences in the remote future.

We believe that the continuity of φ is a natural assumption for g to represent a realistic model, since, if the distance m_k is sampled, then the probability of the event $\{X_{n+1} = +1\}$ should depend smoothly on the magnetization of the interval $[n - m_k + 1, n]$. We will therefore start the analysis assuming only that φ is monotone increasing (which makes g attractive) and satisfies $\varepsilon_* \leq \varphi \leq 1 - \varepsilon_*$, as well as the symmetry condition

$$\varphi(s) + \varphi(-s) = 1. \tag{4}$$

In particular, $\varphi(0) = \frac{1}{2}$. The Bramson–Kalikow mechanism will be developed under these sole assumptions, until additional criteria become necessary to guarantee non-uniqueness. We will give some examples where these criteria are satisfied in Section 6. It is important to notice that all the examples giving non-uniqueness will be functions φ with a rapid increase near $s = 0$: either with a discontinuity (like in the original Bramson–Kalikow example), or with a behavior of the type $\varphi(s) \sim s^{\nu}$, $0 < \nu < 1$. This will be in sharp contrast with the simple example we give in Section 7, of a function φ, smooth at the origin, for which uniqueness holds in a strong sense (for all sequences (p_k), (m_k)). This seems to indicate that a fast increase of φ near zero is a necessary ingredient for non-unicity.

For the Berger–Hoffman–Sidoravicius model, it was shown in Dias and Friedli (2015) that a majority rule φ satisfying a Lipschitz condition in a neighbourhood of the origin always leads to uniqueness.

2 Notation

We identify $\{\pm 1\}$-valued processes with their distributions, that is, with probability measures on $\Omega := \{\pm 1\}^{\mathbb{Z}}$. The elements of Ω are usually denoted by ω. For each k, define the coordinate variable $X_k(\omega) := \omega_k$. When $I \subset \mathbb{Z}$, $\mathcal{F}_I := \sigma(X_k, k \in I)$. If $|I| < \infty$, the elements of $\mathcal{F}_I$ are called cylinders. We let $\mathcal{F} := \mathcal{F}_{\mathbb{Z}}$. A probability measure P on $(\Omega, \mathcal{F})$ is invariant if $P \circ T^{-1} = P$, where $T : \Omega \to \Omega$ is the shift, defined by $(T\omega)_k := \omega_{k+1}$.

We construct invariant probability measures on $(\Omega, \mathcal{F})$ such that the process of the coordinate variables, $X = (X_k)_{k \in \mathbb{Z}}$, is specified by a given g-function g. This is done in the standard way, by taking weak limits. We write Ω as $\{\pm 1\}^{\mathbb{Z}_-} \times \{\pm 1\}^{\mathbb{Z}_+}$, where $\mathbb{Z}_- := \{\dots, -2, -1, 0\}$, $\mathbb{Z}_+ := \{1, 2, \dots\}$. For each $\sigma \in \{\pm 1\}^{\mathbb{Z}_-}$, called boundary condition, we consider the probability measure $P_\sigma := \delta_\sigma \otimes p_\sigma$ on $\{\pm 1\}^{\mathbb{Z}_-} \times \{\pm 1\}^{\mathbb{Z}_+}$, where δ_σ is a Dirac mass at σ, and p_σ is defined, for all cylinder $[a_1^n] := \{X_1 = a_1, \ldots, X_n = a_n\}$, $a_i \in \{\pm 1\}$, by

$$p_\sigma([a_1^n]) := \prod_{k=1}^{n} \{g(a_1^{k-1}\sigma)^{(1+a_k)/2}(1 - g(a_1^{k-1}\sigma))^{(1-a_k)/2}\}, \tag{5}$$

where $a_1^{k-1}\sigma := (a_{k-1}, \ldots, a_1, \sigma_0, \sigma_{-1}, \ldots)$, and $a_1^0\sigma := \sigma$. Clearly, P_σ is not invariant. When $\sigma \equiv +1$, and if g is attractive, then P_σ is denoted simply P_+, and is used to construct an invariant measure $\widehat{P}_+$, as follows (the same can be done when $\sigma \equiv -1$). Let $I \subset \mathbb{Z}$ be finite, and consider a cylinder $C = \{X_i = a_i, i \in I\}$. If n is large enough, then $i + n \in \mathbb{Z}_+$ for all $i \in I$, and thanks to the attractiveness of g, the limit

$$\widehat{P}_+\{C\} := \lim_{n\to\infty} P_+\{T^{-n}C\} \tag{6}$$

can be shown to exist. $\widehat{P}_+$ then extends uniquely to a stationary probability measure on $(\Omega, \mathcal{F})$. It can also be shown that under $\widehat{P}_+$, X is specified by g.

3 The Doeblin–Fortet estimate

This section and the following are devoted to the description of a concentration property satisfied by the measures P_σ, $\widehat{P}_\sigma$, whose associated g-function has summable variations, that is, $\mathrm{var}_k(g) \in \ell^1$:

$$\sum_{k\geq 1} \mathrm{var}_k(g) < \infty. \tag{7}$$

The results presented here will be applied to some auxiliary g-functions that appear in the mechanism of the proof for non-uniqueness, but are of independent interest.

The following result dates back to Doeblin and Fortet's pioneering paper (Doeblin and Fortet, 1937). In essence, it says that when g has summable variation, the limits (6) exist for arbitrary boundary conditions σ, and do not depend on σ. Moreover, the rate of convergence is known explicitly, and is also independent of σ and of the size of the cylinder. The result was stated in this form and proved by Iosifescu (1992).

Theorem 3.1. *Assume g is regular and $\mathrm{var}_k(g) \in \ell^1$. Then there exists a unique invariant measure $\widehat{P}$ on $(\Omega, \mathcal{F})$ such that for all cylinder A, $\widehat{P}\{A\} = \lim_{n\to\infty} P_\sigma\{T^{-n}A\}$, uniformly in σ. More precisely,*

$$\left|\widehat{P}\{A\} - P_\sigma\{T^{-n}A\}\right| \leq \phi(n) \qquad \forall n \geq 1, \tag{8}$$

where the $\phi(n)$ are mixing coefficients given by

$$\phi(n) := \inf_{1\leq m\leq n}\left\{2\left(1 - (1-\varepsilon_*)^m\right)^{n/m-1} + 2\sum_{k\geq m} \mathrm{var}_k(g)\right\}. \tag{9}$$

Observe that $\phi(n)$ depends on g only through ε_* and $\mathrm{var}_k(g)$, and that $\phi(n) \to 0$ when $n \to \infty$ since g has summable variation. Theorem 3.1 implies the following mixing property for $\widehat{P}$:

Corollary 3.1. *For all $k \in \mathbb{Z}$ and $j \geq 1$, if $A \in \mathcal{F}_{(-\infty,k]}$ and $B \in \mathcal{F}_{[k+j,+\infty)}$ are two cylinders, then*

$$\left|\widehat{P}\{B \cap A\} - \widehat{P}\{B\}\widehat{P}\{A\}\right| \leq \phi(j)\widehat{P}\{A\}. \tag{10}$$

Proof. Namely, assume A is a cylinder of the form $A = [a_i^k] \in \mathcal{F}_{[i,k]}$, $i \leq k$, and $B \in \mathcal{F}_{[k+j,+\infty)}$. By (8) we have, for any $\sigma \in \{\pm 1\}^{\mathbb{Z}_-}$,

$$\begin{aligned}\widehat{P}\{B \cap A\} &= \lim_{n\to\infty} P_\sigma\{T^{-n}B \cap T^{-n}A\} \\ &= \lim_{n\to\infty} \sum P_\sigma\{T^{-n}B \cap T^{-n}A \cap [\omega_1^{i+n-1}]\},\end{aligned}$$

where the sum is over all configurations $\omega_1^{i+n-1} = (\omega_1, \ldots, \omega_{i+n-1})$ on $[1, i + n - 1]$. Let $\tilde{\sigma} = (a_k, \ldots, a_i, \omega_{i+n-1}, \ldots, \omega_1, \sigma_0, \sigma_{-1}, \ldots)$. Then by the definition of $P_{\tilde{\sigma}}$, (8) and the invariance of $\widehat{P}$,

$$\begin{aligned}P_\sigma\{T^{-n}B | T^{-n}A \cap [\omega_1^{i+n-1}]\} &= P_{\tilde{\sigma}}\{T^k B\} \\ &\leq \widehat{P}\{T^k B\} + \phi(j) = \widehat{P}\{B\} + \phi(j).\end{aligned}$$

Therefore, $\widehat{P}\{B \cap A\} \leq (\widehat{P}\{B\} + \phi(j))\widehat{P}\{A\}$. □

4 A concentration inequality

Under the same summability hypothesis (7) on g, our aim here is to obtain a concentration inequality for the empirical magnetization $(X_1 + \cdots + X_n)/n$, valid for *all* n, first under $\widehat{P}$, and then under some P_σ.

The main ingredient is a result of Samson (2000). Let $(Y_1, Y_2, \ldots, Y_n)$ be a sample of bounded random variables $0 \leq Y_i \leq 1$ defined on some common probability space. The dependencies among the variables Y_i are measured through an $n \times n$ triangular matrix Γ whose elements Γ_{ij} are defined as follows. For $1 \leq k \leq l \leq n$, let $Y_k^l := (Y_k, \ldots, Y_l)$. Let also $\mathcal{L}(Y_j^n | Y_1^{i-1} = a_1^{i-1}, Y_i = c_i)$ denote the distribution of Y_j^n conditioned on $\{Y_1 = a_1, \ldots, Y_{i-1} = a_{i-1}, Y_i = c_i\}$. Then, define $\Gamma_{ij} := 0$ if $i > j$, $\Gamma_{ii} := 1$, and for $i < j$,

$$\begin{aligned}(\Gamma_{ij})^2 := \sup_{a_1^{i-1}, b_1^{i-1}, c_i} &\big\|\mathcal{L}(Y_j^n | Y_1^{i-1} = a_1^{i-1}, Y_i = c_i) \\ &- \mathcal{L}(Y_j^n | Y_1^{i-1} = b_1^{i-1}, Y_i = c_i)\big\|_{\mathrm{TV}}.\end{aligned}$$

Here, $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation of signed measures. By Corollary 4 in Samson (2000), for every function $f = f(Y_1, \dots, Y_n)$, convex and l-Lipschitz for some $l \le 1$,

$$P\{|f - E[f]| \ge t\} \le 2\exp\Big(-\frac{t^2}{2\|\Gamma\|^2}\Big) \qquad \forall t \ge 0, \tag{11}$$

where $\|\Gamma\|$ is the operator norm of Γ with respect to the Euclidean metric on $\mathbb{R}^n$. Equation (11) is useful when $\|\Gamma\|$ is bounded uniformly in the size of the sample.

Theorem 4.1. *Let g be regular with summable variation and let $\phi(n)$ be defined as in (9). Define $\gamma = \gamma(g)$ by*

$$\gamma := 1 + \sum_{n\ge 1} \sqrt{\phi(n)}. \tag{12}$$

If $\gamma < \infty$, then for all $0 < \delta < 1$,

$$\widehat{P}\Big\{\Big|\frac{X_1 + \cdots + X_n}{n} - \widehat{E}[X_1]\Big| \ge \delta\Big\} \le 2\exp\Big(-\frac{\delta^2}{16\gamma^2} n\Big) \qquad \forall n \ge 1. \tag{13}$$

Proof. Let $Y_j := \frac{X_j+1}{2}$, that is, $Y_j \in \{0, 1\}$. We apply (11) with $P := \widehat{P}$, with the 1-Lipschitz function $f(x_1, \dots, x_n) := n^{-1/2}(x_1 + \cdots + x_n)$, and $t = \frac{\delta}{2}\sqrt{n}$. The invariance of $\widehat{P}$ gives $\widehat{E}[Y_1] = \widehat{E}[f]/\sqrt{n}$, and so

$$\begin{aligned}
\widehat{P}\Big\{\Big|\frac{X_1 + \cdots + X_n}{n} - \widehat{E}[X_1]\Big| \ge \delta\Big\} &= \widehat{P}\Big\{\Big|\frac{Y_1 + \cdots + Y_n}{n} - \widehat{E}[Y_1]\Big| \ge \frac{\delta}{2}\Big\} \\
&= \widehat{P}\Big\{|f - \widehat{E}[f]| \ge \frac{\delta}{2}\sqrt{n}\Big\} \\
&\le 2\exp\Big(-\frac{\delta^2}{8\|\Gamma\|^2} n\Big).
\end{aligned}$$

Then, observe that for all $i < j$ (the second inequality follows by (10))

$$(\Gamma_{ij})^2 \le \sup_{A, A' \in \mathcal{F}_{[1,i]}} \sup_{B \in \mathcal{F}_{[j,n]}} |\widehat{P}\{B|A\} - \widehat{P}\{B|A'\}| \le 2\phi(j-i). \tag{14}$$

Now for all $x \in \mathbb{R}^n$, Γx can be written as

$$\Gamma x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} + \begin{pmatrix} \Gamma_{12}x_2 \\ \Gamma_{23}x_3 \\ \vdots \\ \Gamma_{n-1,n}x_n \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} \Gamma_{1,n-1}x_{n-1} \\ \Gamma_{2n}x_n \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \Gamma_{1n}x_n \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

Together with (14), this implies that

$$\|\Gamma x\| \le \|x\| + \sqrt{2\phi(1)}\|x\| + \cdots + \sqrt{2\phi(n-1)}\|x\| \le \sqrt{2}\gamma\|x\|,$$

which gives $\|\Gamma\| \le \sqrt{2}\gamma$, and proves (13). $\square$

As a corollary of Theorems 3.1 and 4.1, we obtain a concentration inequality for P_σ, uniform in σ.

Theorem 4.2. *Let g be regular with $\mathrm{var}_k(g) \in \ell^1$, and assume $\gamma = \gamma(g) < \infty$. Let $0 < \delta < 1$. Then there exists a numerical constant $c > 0$ such that for all $n \geq 1$ and all integer $l \leq \frac{\delta}{4}n$, uniformly in σ,*

$$P_\sigma\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \widehat{E}[X_1]\right| \geq \delta\right\} \leq 2\exp\left(-c\frac{\delta^2}{\gamma^2}n\right) + \phi(l). \tag{15}$$

Proof. If $l \geq 1$ is such that $\frac{l}{n} \leq \frac{\delta}{4}$, then

$$\begin{aligned}
&P_\sigma\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \widehat{E}[X_1]\right| \geq \delta\right\} \\
&\quad\leq P_\sigma\left\{\left|\frac{X_{l+1} + \cdots + X_n}{n-l} - \widehat{E}[X_1]\right| \geq \frac{\delta}{2}\right\} \\
&\quad= P_\sigma\left\{T^{-l}\left\{\left|\frac{X_1 + \cdots + X_{n-l}}{n-l} - \widehat{E}[X_1]\right| \geq \frac{\delta}{2}\right\}\right\} \\
&\quad\leq \widehat{P}\left\{\left|\frac{X_1 + \cdots + X_{n-l}}{n-l} - \widehat{E}[X_1]\right| \geq \frac{\delta}{2}\right\} + \phi(l),
\end{aligned}$$

and using (13), (15) holds with $c := \frac{3}{256}$. □

5 The Bramson–Kalikow mechanism

In this section, we follow Bramson and Kalikow, and give two sufficient conditions on a g-function of the form (2) that guarantee the existence of at least two distinct processes specified by g. We will first express these conditions in a general form, and in Section 6 consider examples.

The reader familiar with Bramson and Kalikow (1993) can directly check how estimates (13) and (14) in Bramson and Kalikow (1993), obtained with the ergodic theorem for Markov chains, will be replaced by the more explicit concentration inequality (15).

Non-uniqueness is obtained by showing that

$$\widehat{P}_+\{X_0 = +1\} > \frac{1}{2} > \widehat{P}_-\{X_0 = +1\}. \tag{16}$$

To obtain (16), we study the magnetization inside large blocks in the past. For all $i \in \mathbb{Z}$, $k \geq 1$, define $B_k(i) := [i - m_k, i) \cap \mathbb{Z}$, $B^k(i) := (i, i + m_k] \cap \mathbb{Z}$, which both contain m_k points of $\mathbb{Z}$. For the first inequality in (16), we will show that all blocks are positively magnetized with high P_+-probability. For a given $0 < \delta_k < 1$, say that $B_k(i)$ is good if

$$\frac{1}{m_k} \sum_{j \in B_k(i)} X_j \geq \delta_k, \tag{17}$$

and bad otherwise; the same can be defined for the block $B^k(i)$. The point will be to find a decreasing sequence $\delta_k \searrow 0$, such that for all $i \in \mathbb{Z}_+$,

$$P_+\{B_k(i) \text{ good}\} \geq 1 - 3^{-k}\delta_k \qquad \forall k \geq 1. \tag{18}$$

Namely, if we assume for a while that (18) holds, one can fix $k \geq 1$ and compute, for any large i,

$$E_+\left[\frac{1}{m_k} \sum_{j \in B_k(i)} X_j\right] \geq \delta_k P_+\{B_k(i) \text{ good}\} - P_+\{B_k(i) \text{ bad}\} \geq 3^{-1}\delta_k.$$

Therefore,

$$\frac{1}{m_k} \sum_{j \in B_k(i)} P_+\{X_j = +1\} = \frac{1}{2} + \frac{1}{2} E_+\left[\frac{1}{m_k} \sum_{j \in B_k(i)} X_i\right] \geq \frac{1}{2}(1 + 3^{-1}\delta_k).$$

By taking $i \to \infty$, and using (6), we get

$$\widehat{P}_+\{X_0 = +1\} = \lim_{j \to \infty} P_+\{X_j = +1\} \geq \frac{1}{2}(1 + 3^{-1}\delta_k) > \frac{1}{2}.$$

This lower bound holds for all k, and it is of course optimal by taking the smallest possible k, which is $k = 1$.

The two following sections are devoted to finding conditions under which (18) holds.

5.1 The induction step

Equation (18) is shown by induction on $i \geq 1$. If $i = 1$, then (18) holds (by the definition of P_+). Assume then that (18) has been proved for all i with $i < i_0$. We fix some $i < i_0$, some $k \geq 1$, and study $P_+\{B^k(i) \text{ good}\} = 1 - P_+\{B^k(i) \text{ bad}\}$. By the attractiveness of the model, the presence of good blocks in the past of i favors a positive magnetization in the near future of i. We thus use the induction hypothesis and condition $\{B^k(i) \text{ bad}\}$ on an event in the past that favorizes the positivity of the magnetization in $B^k(i)$, and that by the induction step has a sufficiently large probability. Let therefore (our notation differs slightly from Bramson and Kalikow, 1993)

$$G_k(i) := \bigcap_{j > k} \{B_j(i) \text{ good}\}.$$

Then,

$$P_+\{B^k(i) \text{ bad}\} \leq P_+\{B^k(i) \text{ bad} | G_k(i)\} + P_+\{G_k(i)^c\}. \tag{19}$$

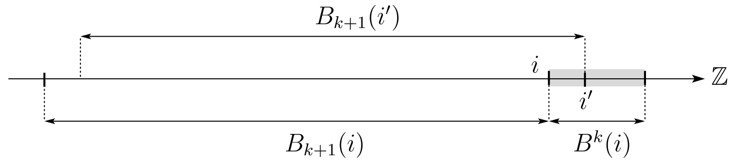

Figure 1 *The usefulness of lacunarity: if $B_{k+1}(i)$ is positively magnetized and $m_k \ll m_{k+1}$, then $B_{k+1}(i')$ is also positively magnetized.*

By the induction hypothesis, and since δ_k is decreasing,

$$P_+\{G_k(i)^c\} \le \sum_{j>k} P_+\{B_j(i) \text{ bad}\} \le \sum_{j>k} 3^{-j}\delta_j \le \frac{1}{2}3^{-k}\delta_k. \tag{20}$$

The point is then to show that

$$P_+\{B^k(i) \text{ bad}|G_k(i)\} \le \frac{1}{2}3^{-k}\delta_k. \tag{21}$$

One way of obtaining (21) is to assume that the sequence m_k is lacunary. Namely, observe that the blocks that appear in the conditioning event $G_k(i)$ are all of size at least m_{k+1} (see Figure 1). Therefore, if m_{k+1} is much larger than m_k, then the goodness of each block $B_j(i)$, $j > k$, implies that for any $i' \in B^k(i)$, the magnetization inside $B_j(i')$ is always close to the magnetization of $B_j(i)$, which is strictly positive. This allows to extract from g a contribution which is constant over $B^k(i)$, which we call magnetic field, and allows to reduce the study of the process on $B^k(i)$ to that of a simpler effective model. The point is then to see under which conditions this field is positive, and to quantify the way in which it creates a positive magnetization inside $B^k(i)$. This is done in the following section.

5.2 The reduction

We define the g-function at a point $i \in \mathbb{Z}$ by

$$g_i := g(X_{i-1}, X_{i-2}, \dots) = \sum_{n\ge 1} p_n \varphi\big(b_n(i)\big),$$

where $b_n(i) := \frac{1}{m_n}\sum_{j\in B_n(i)} X_j$ is the magnetization of $B_n(i)$.

Lemma 5.1. *Let $i' \in B^k(i)$. Then on the event $G_k(i)$,*

$$g_{i'} \ge q_{i'}^{(k-1)} + h_k, \tag{22}$$

where

$$q_{i'}^{(k-1)} := \sum_{n=1}^{k-1} p_n \varphi\big(b_n(i')\big) + \frac{1}{2}\sum_{n\ge k} p_n, \tag{23}$$

and where h_k is a constant magnetic field defined by

$$h_k := \sum_{n>k} p_n \left\{ \varphi(\delta_n - 2m_k/m_n) - \frac{1}{2} \right\} - \left(\frac{1}{2} - \varepsilon_* \right) p_k. \tag{24}$$

Proof. Write $g_{i'} = \sum_{n=1}^{k-1} p_n \varphi(b_n(i')) + p_k \varphi(b_k(i')) + \sum_{n>k} p_n \varphi(b_n(i'))$, and observe that on $G_k(i)$ we have, when $n > k$,

$$b_n(i') \geq b_n(i) - \frac{2|i'-i|}{m_n} \geq \delta_n - \frac{2|i'-i|}{m_n} \geq \delta_n - \frac{2m_k}{m_n}.$$

When $n = k$ we use $b_k(i') \geq -1$. The monotonicity of φ thus implies (22). $\square$

We have used the notation $q^{(k-1)}$ to indicate that the dependence of this g-function on the past is only a distance m_{k-1} back:

$$\mathrm{var}_j(q^{(k-1)}) = 0 \qquad \text{if } j > m_{k-1}. \tag{25}$$

By Theorem 3.1 and the symmetry (4), any process specified by $q^{(k-1)}$ has zero magnetization. We will of course aim at using the lower bound (22) on $g_{i'}$ when $h_k > 0$.

Since the boundary condition is fixed, the event $G_k(i)$ can be identified with a set of configurations on the interval $[1, i]$, which we denote by $\sigma_1^i = (\sigma_1, \dots, \sigma_i)$. Therefore,

$$P_+\{B^k(i) \text{ bad} | G_k(i)\} \leq \sup_{\sigma_1^i \in G_k(i)} P_+\{B^k(i) \text{ bad} | \sigma_1^i\}.$$

By Lemma 5.1, and since $\{B^k(i) \text{ bad}\}$ is a decreasing event, a standard coupling gives

$$P_+\{B^k(i) \text{ bad} | \sigma_1^i\} \leq P^{(k)}_{\sigma_1^i +}\{B^k(0) \text{ bad}\}, \tag{26}$$

where $P^{(k)}_{\sigma_1^i +}$ is constructed as in (5) with the g-function $q^{(k-1)} + h_k$ and the boundary condition $\sigma_1^i + := (\sigma_i, \dots, \sigma_1, +, +, \dots) \in \{\pm 1\}^{\mathbb{N}}$. Since $q^{(k-1)} + h_k$ has summable variation, we can consider the associated invariant measure $\widehat{P}^{(k)}$ constructed in Theorem 3.1. Let $\widehat{E}^{(k)}$ denote the expectation under $\widehat{P}^{(k)}$. If $h_k \geq 0$, then $\widehat{P}^{(k)}$ can further be coupled to the measure specified by $q^{(k-1)}$, yielding $\widehat{E}^{(k)}[X_1] \geq h_k$. In order for $B^k(i)$ to be good, the magnetic field must be larger than δ_k. We thus assume that

$$h_k \geq 2\delta_k, \tag{27}$$

and get

$$\begin{aligned} P^{(k)}_{\sigma_1^i +}\{B^k(0) \text{ bad}\} &= P^{(k)}_{\sigma_1^i +}\left\{ \frac{X_1 + \cdots + X_{m_k}}{m_k} < \delta_k \right\} \\ &\leq P^{(k)}_{\sigma_1^i +}\left\{ \left| \frac{X_1 + \cdots + X_{m_k}}{m_k} - \widehat{E}^{(k)}[X_1] \right| \geq \delta_k \right\}. \end{aligned} \tag{28}$$

Let $\phi_{k-1}(n)$, $n \geq 1$, denote the mixing coefficients (9) associated to $q^{(k-1)} + h_k$, and γ_{k-1} denote the series in (12), with ϕ_{k-1} in place of ϕ:

$$\gamma_{k-1} := 1 + \sum_{n \geq 1} \sqrt{\phi_{k-1}(n)}. \tag{29}$$

Using Theorem 4.2 for (28), with $l := \lfloor \frac{\delta_k}{4} m_k \rfloor$, we get

$$P_+\{B^k(i) \text{ bad} | G_k(i)\} \leq 2\exp\Big(-c\frac{\delta_k^2}{\gamma_{k-1}^2} m_k\Big) + \phi_{k-1}\Big(\Big\lfloor \frac{\delta_k}{4} m_k \Big\rfloor\Big). \tag{30}$$

One thus sees that (21) holds if

$$2\exp\Big(-c\frac{\delta_k^2}{\gamma_{k-1}^2} m_k\Big) + \phi_{k-1}\Big(\Big\lfloor \frac{\delta_k}{4} m_k \Big\rfloor\Big) \leq \frac{1}{2} 3^{-k} \delta_k, \tag{31}$$

which is clearly satisfied when m_k is large enough.

So far, we have proved the following theorem.

Theorem 5.1. *Let g be of the form (2), where φ is non-decreasing and satisfies $\varphi(s) + \varphi(-s) = 1$. Let (δ_k) be a decreasing sequence $\delta_k > 0$. If (p_k), (m_k), φ and (δ_k) satisfy (27) and (31) for large enough k, then*

$$\widehat{P}_+\{X_0 = +1\} > \frac{1}{2} > \widehat{P}_-\{X_0 = +1\}.$$

In particular, g specifies at least two distinct stationary processes.

A look at the proof above shows that a slightly more general statement holds. Call a boundary condition σ of type $+$ (resp. $-$) if

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \sigma_i > 0, \qquad \Big(\text{resp. } \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \sigma_i < 0\Big).$$

For example, the boundary condition defined by $\sigma_i := -1$ if $1 \leq i \leq L$, $\sigma_i := +1$ if $i > L$, is of type $+$. It can be shown that if σ^+ is of type $+$ and σ^- is of type $-$, then under the same hypothesis as in Theorem 5.1,

$$\liminf_{n\to\infty} P_{\sigma^+}\{X_n = +1\} > \frac{1}{2} > \limsup_{n\to\infty} P_{\sigma^-}\{X_n = +1\}. \tag{32}$$

6 Examples

Since the conditions (27) and (31) combine the sequences (p_n), (m_n), (δ_n) and the function φ in a rather intricate way, the search for concrete examples is a delicate task.

To start, consider (27). In order for h_k (defined in (24)) to be non-negative, we first require that $\varphi(\delta_n - 2m_k/m_n) \geq \frac{1}{2}$. This can be done by assuming that

$$\frac{m_{k-1}}{m_k} \leq \frac{\delta_k}{4} \qquad \forall k \geq 2. \tag{33}$$

When this holds, then

$$h_k \geq \sum_{n>k} p_n \left\{ \varphi(\delta_n/2) - \frac{1}{2} \right\} - \left(\frac{1}{2} - \varepsilon_* \right) p_k. \tag{34}$$

Below, we will first choose (p_n) and φ, and find a proper sequence (δ_n) that satisfies (27). To complete the construction of the g-function, (m_k) can then always be taken so as to satisfy (31) and (33). We will not bother choosing (m_k) in an optimal manner, that is, taking the smallest m_k realizing simultaneously (31) and (33). This would mean searching for non-uniqueness with the least lacunary sequence, and with the fastest possible decaying variation. Unfortunately, as a tedious but straightforward computation shows, in any of the examples below, m_k always grows at least as fast as $k!$, and the variation obtained never belongs to any ℓ^p, even for p large.

6.1 The original Bramson–Kalikow example

Consider the pure majority rule φ_{PMR} defined in (3) and represented on Figure 2, and assume that m_k is odd for all k. Then $\varphi_{\mathrm{PMR}}(\delta_n/2) - \frac{1}{2} = \frac{1}{2} - \varepsilon_*$. Assume further that (p_n) is such that $\sum_{n>k} p_n \geq 2p_k$ for all large enough k. Bramson and Kalikow mention $p_n = cr^n$ with $r \in (\frac{2}{3}, 1)$, $c = (1-r)/r$, but any sequence[2] $p_n \sim \frac{1}{n^{1+\varepsilon}}$ with $0 < \varepsilon < 1$ also satisfies this condition. Then, (34) becomes $h_k \geq (\frac{1}{2} - \varepsilon_*)p_k$. Therefore, δ_k can simply be chosen as $2\delta_k \equiv (\frac{1}{2} - \varepsilon_*)p_k$, and (27) holds.

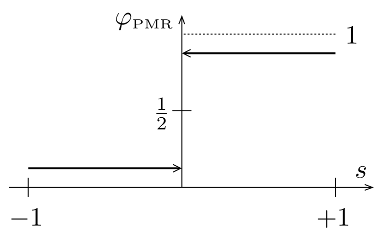

Figure 2 *The pure majority rule used in Bramson and Kalikow* (1993).

[2] $a_n \sim b_n$ means there exists two positive constants c_-, c_+ such that $c_- \leq a_n/b_n \leq c_+$ for all large enough n.

6.2 Some g-functions with φ continuous at $s = 0$

Assume φ is continuous at $s = 0$, and that it satisfies, for small s,

$$\left|\varphi(s) - \frac{1}{2}\right| \geq cs^{\nu}, \tag{35}$$

for some constant $c > 0$ and some $0 < \nu < 1$ (it is easy to see that if $\nu \geq 1$, then (27) is never satisfied). Then for large k, (34) becomes

$$h_k \geq c/2^{\nu} \sum_{n>k} p_n \delta_n^{\nu} - \left(\frac{1}{2} - \varepsilon_*\right) p_k. \tag{36}$$

Take

$$p_n \sim \frac{1}{n^{1+\varepsilon}} \qquad \text{with } 0 < \varepsilon < 1.$$

Then it is easy to verify that by taking

$$\delta_n := \frac{1}{n^{\tau}} \qquad \text{with } 0 < \tau < 1/\nu,$$

the lower bound in (36) is larger than $2\delta_k$ for large k, thus fulfilling (27). Taking

$$p_n \sim \frac{1}{n(\log n)^{1+\varepsilon}} \qquad \text{with } 0 < \varepsilon < 1,$$

with the same δ_n, also ensures that (27) is satisfied for large k.

7 On the role played by $\varphi'(0)$

A common feature of the examples above is that in each of them, φ increases fast near $s = 0$. Loosely speaking: $\varphi'(0) = \infty$. A natural question is to know if this is a necessary ingredient for non-uniqueness. The following example illustrates a case where the finiteness of $\varphi'(0)$ leads to uniqueness among a large class of g-functions (of the form (2)).

Let $0 < \varepsilon_* < 1/2$ and set (see Figure 3)

$$\varphi_*(s) := \frac{1}{2} + \left(\frac{1}{2} - \varepsilon_*\right) s. \tag{37}$$

Theorem 7.1. *Let φ_* be as above, choose any sequences (p_n), (m_n), and denote the associated g-function by g_*. Then, g_* specifies a unique*[3] *stationary process.*

[3]Uniqueness should be here understood in the sense that all processes specified by g_* are identically distributed.

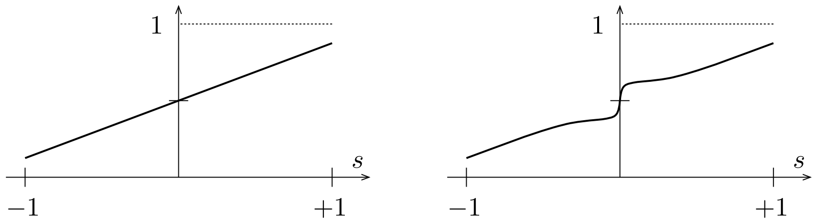

Figure 3 *On the left, φ_*, which specifies a unique process regardless of the sequences (p_k) and (m_k). On the right, an arbitrarily small perturbation of φ_* for which non-uniqueness holds.*

Proof. Let $X = (X_n)_{n\in\mathbb{Z}}$ be any stationary process specified by g_*. Let

$$m := E[X_0]. \tag{38}$$

We have $E[X_0] = 2P\{X_0 = +1\} - 1$, and

$$P\{X_0 = +1\} = \int P\{X_0 = +1|X_{-1}, X_{-2}, \ldots\}\, dP = E[g_*].$$

Due to the form of φ_* and to the stationarity of X,

$$E[g_*] = \sum_{n\geq 1} p_n E\left[\varphi_*\left(\frac{1}{m_n}\sum_{i=1}^{m_n} X_{-i}\right)\right] = \varphi_*\big(E[X_0]\big) = \varphi_*(m).$$

Therefore, m is solution of the fixed-point equation

$$\frac{1+m}{2} = \varphi_*(m), \tag{39}$$

whose unique solution, since $0 < \varepsilon_* < \frac{1}{2}$, is $m = 0$. Considering in particular the processes X^+ and X^- prepared with the boundary condition $+$, respectively $-$ (i.e., the coordinate processes associated to $\widehat{P}_+$, resp. $\widehat{P}_-$), the above shows that $E[X_0^+] = E[X_0^-]$. As shown by Hulse (1991), this implies that, due to the attractiveness of g_*, X^+ and X^- have the same distribution, and that any stationary process specified by g_* has the same distribution as X^+ (and X^-). □

Since the above argument does not depend on the sequences (p_n) and (m_n), one can construct examples where uniqueness holds with slowly decaying variations.[4] Namely, observe first that with φ_* defined as in (37),

$$\mathrm{var}_k(g_*) = (1 - 2\varepsilon_*) \sum_{n>n(k)} p_n\left(1 - \frac{k}{m_n}\right), \tag{40}$$

[4] Similar features had already been observed by Hulse (1991).

where $n(k)$ is the largest integer n for which $m_n \le k$. We consider some particular cases. In the non-lacunary case ($m_n := n$), with $p_n \sim \frac{1}{n^{1+\varepsilon}}$, (40) gives

$$\mathrm{var}_k(g_*) \sim \frac{1}{k^\varepsilon},$$

that is, $\mathrm{var}_k(g_*) \in \ell^p$ if and only if $p > 1/\varepsilon$. In particular, if $\varepsilon < 1/2$, then $\mathrm{var}_k(g_*) \notin \ell^2$, giving uniqueness in a case not covered by the Johansson–Öberg criterion. On the other hand, if m_n grows faster than n, for example when it satisfies the lacunarity condition

$$\liminf_{n\to\infty} \frac{m_n}{n} > 1, \tag{41}$$

then for large k, since $n(k) \le k$,

$$\mathrm{var}_k(g_*) \ge c_1 \sum_{n>k} p_n$$

for some $c_1 > 0$. If one considers for example, $p_n \sim n^{-1}(\log n)^{-1-\varepsilon}$, then for large k,

$$\mathrm{var}_k(g_*) \ge \frac{c_1}{2\varepsilon(\log k)^\varepsilon}.$$

This gives an example of an attractive g-function whose variation of order k is not in any ℓ^p, but which specifies a unique stationary process.

These examples of uniqueness are in sharp contrast with those of Section 6. In particular, we see that it is possible to choose a pair of sequences (p_n), (m_n) (typically very lacunary, satisfying the requirements of Section 6.2) for which using φ_* leads to uniqueness, but in which an arbitrarily small perturbation of φ_* (see the right of Figure 3) with $\varphi'(0) = \infty$, leads to non-uniqueness.

This discussion raises the question of knowing if existence and finiteness of $\varphi'(0)$ always implies uniqueness, or whether the existence of non-trivial solutions to fixed-point equations of the form (39) (as happens in Bramson and Kalikow's original example, or when $\varphi'(0) = +\infty$), combined with well chosen sequences (p_n) and (m_n), can lead to non-uniqueness. This is work in progress.

In Dias and Friedli (2015), we gave a closer look at the model of Berger–Hoffman–Sidoravicius, which also involves a majority rule φ. Besides providing a deeper analysis of the model, and a complete study of the uniqueness and non-uniqueness regimes, we also showed that in general, a local Lipschitz condition on φ leads to uniqueness.

Acknowledgments

I am grateful to S. Gallo, M. E. Vares and V. Sidoravicius for many useful discussions and remarks. This work benefited of partial financial support from CNPq (Bolsa de Produtividade), and from IMPA (Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro).

References

Berger, N., Hoffman, C. and Sidoravicius, V. (2003). Non-uniqueness for specifications in $\ell^{2+\varepsilon}$. Available at arXiv:math/0312344.

Bramson, M. and Kalikow, S. (1993). Non-uniqueness in g-functions. *Israel J. Math.* **84**, 153–160. MR1244665

Dias, J. C. A. and Friedli, S. (2015). Uniqueness vs. non-uniqueness in complete connections with modified majority rules. *Probab. Theory Related Fields*. To appear.

Doeblin, W. and Fortet, R. (1937). Sur les chaînes à liaisons complètes. *Bull. Soc. Math. France* **65**, 132–148. MR1505076

Hulse, P. (1991). Uniqueness and ergodic properties of attractive g-measures. *Ergodic Theory Dynam. Systems* **11**, 65–77. MR1101084

Iosifescu, M. (1992). A coupling method in the theory of dependence with complete connections according to Doeblin. *Rev. Roumaine Math. Pures Appl.* **37**, 59–66. MR1172269

Johansson, A. and Öberg, A. (2003). Square summability of variations of g-functions and uniqueness of g-measures. *Math. Res. Lett.* **10**, 587–601. MR2024717

Samson, P.-M. (2000). Concentration of measure inequalities for Markov chains and ϕ-mixing processes. *Ann. Probab.* **28**, 416–461. MR1756011

Departamento de Matemática, ICEx
Universidade Federal de Minas Gerais
Av. Antônio Carlos 6627, C.P. 702
CEP 30123-970, Belo Horizonte
Brasil
E-mail: sacha@mat.ufmg.br