Uniqueness vs. non-uniqueness for complete connections with modified majority rules

J. C. A. Dias[1,2] · S. Friedli[3]

Abstract We take a closer look at a class of chains with complete connections introduced by Berger, Hoffman and Sidoravicius. In particular, besides giving a sharper description of the uniqueness and non-uniqueness regimes, we show that if the pure majority rule used to fix the dependence on the past is replaced with a function that is Lipschitz at the origin, then uniqueness always holds, even with arbitrarily slow decaying variation.

Until January 2014, after this paper was submitted, J. C. A. Dias was member of the Departamento de Matemática, at the Universidade Federal de Ouro Preto.

✉ S. Friedli
  sacha@mat.ufmg.br

[1] Departamento de Física e Matemática, CEFET-MG, Av. Amazonas 7675,
    Belo Horizonte CEP 30510-000, Brazil

[2] Present Address: Departamento de Matemática, Universidade Federal de Minas Gerais,
    Av. Antônio Carlos 6627, C.P. 702, Belo Horizonte CEP 30123-970, Brazil

[3] Departamento de Matemática, Universidade Federal de Minas Gerais,
    Av. Antônio Carlos 6627, C.P. 702, Belo Horizonte CEP 30123-970, Brazil

1 Introduction

We consider stationary stochastic processes on $\mathbb{Z}$,

$$\ldots, Z_{-2}, Z_{-1}, Z_0, Z_1, Z_2, \ldots$$

where each Z_t, $t \in \mathbb{Z}$, is a symbol taking values in a finite alphabet A. The processes we consider are called chains with complete connections (Doeblin and Fortet [3]), due to a dependence on the past of the following form. Assume some measurable map $g : \mathsf{A} \times \mathsf{A}^{\mathbb{N}} \to [0, 1]$ is given a priori, called g-function, and that for all t, all $z_t \in \mathsf{A}$,

$$P(Z_t = z_t | Z_{t-1} = z_{t-1}, Z_{t-2} = z_{t-2}, \ldots) = g(z_t | z_{t-1}, z_{t-2}, \ldots) \quad \text{a.s.} \quad (1)$$

A processes $Z = (Z_t)_{t \in \mathbb{Z}}$ satisfying (1) is said to be specified by g. The role played by g for Z is therefore analogous to a transition kernel for a discrete time Markov process, except that it allows dependencies on the whole past of the process.

We will always assume that g is regular, which means that it satisfies the following two conditions.

(1) It is uniformly bounded away from 0 and 1: there exists $\eta > 0$ such that $\eta \leq g(z_0 | z) \leq 1 - \eta$ for all $z_0 \in \mathsf{A}$, $z \in \mathsf{A}^{\mathbb{N}}$.
(2) Define the variation of g of order j by

$$\mathrm{var}_j(g) := \sup |g(z_0 | z) - g(z_0 | z')|,$$

where the sup is over all $z_0 \in \mathsf{A}$, and over all $z, z' \in \mathsf{A}^{\mathbb{N}}$ for which $z_i = z_i'$ for all $1 \leq i \leq j$. Then g is continuous in the sense that $\mathrm{var}_j(g) \to 0$ when $j \to \infty$.

When g is regular, the existence of at least one stationary process specified by g follows by a standard compactness argument (see also the explicit construction given below). Once existence is guaranteed, uniqueness can be shown under additional assumptions on the speed at which $\mathrm{var}_j(g) \to 0$. For instance, Doeblin and Fortet [3] showed that if

$$\sum_j \mathrm{var}_j(g) < \infty,$$

then there exists a unique process specified by g. More recently, Johansson and Öberg [9] strengthned this result, showing that uniqueness holds as soon as

$$\sum_j \mathrm{var}_j(g)^2 < \infty. \quad (2)$$

An interesting and natural question is to determine if a given regular g-function can lead to a phase transition, that is if it specifies at least two *distinct* processes.

In a pioneering paper, Bramson and Kalikow [2] gave the first example of a regular g-function exhibiting a phase transition. More recently, Berger et al. [1], in a remarkable paper, introduced a model whose g-function also exhibits a phase transition, but whose
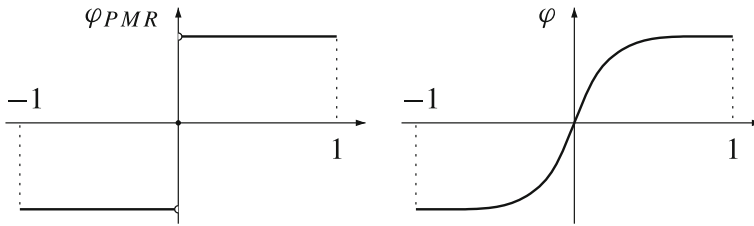
Fig. 1 On the *left*, the pure majority rule used in [1], for which non-uniqueness holds when $0 < \alpha < \frac{1-\epsilon_*}{2}$. On the *right*, a smoothed version, for which the process is always unique for all $\alpha > 0$, or more generally, for all sequence $h_k \searrow 0$

variation has a summability that can be made arbitrarily close to the ℓ^2-summability of the Johansson–Öberg criterion (see Remark 2 below).

The *g*-functions constructed in [2] and [1] have common features. The main one is that they both rely on some *majority rule* used in order to fix the influence of the past on the probability distribution of the present. That is, Z_{t+1}, given $(Z_s)_{s \leq t}$, is determined by the *sign* (and not the true value) of the average of a subset of the variables $(Z_s)_{s \leq t}$ over a large finite region. This feature is essential in the mechanisms that lead to non-uniqueness, since it allows, roughly speaking, small local fluctuations to have dramatic effects in the remote future, thus favoring the transmission of information from $-\infty$ to $+\infty$.

For the Bramson–Kalikow model, it had already been observed in [5] that arbitrarily small changes in the behavior of the majority rule, turning it smooth at the origin, can have important consequences on uniqueness/non-uniqueness of the process.

In this paper, we give a closer look at a class of models based on the one of Berger, Hoffman, and Sidoravicius (which will be called simply the BHS-model hereafter). Beyond giving a more detailed description of the original model of [1], our results show that any smoothing of the majority rule (see Fig. 1) leads, under general assumptions, to uniqueness, even for very slow-decaying variations.

We will present these models from scratch, and not assume any prior knowledge concerning [1]; since their construction is not trivial and deserves some explanations, we will state our results precisely only at the end of Sect. 2.

Before proceeding, we single out other non-uniqueness-related works. Hulse [8] gave examples of non-uniqueness, based on the Bramson–Kalikow approach. Fernández and Maillard [4] constructed an example, using a long-range spin system of statistical mechanics, although in a non-shift-invariant framework. Gallesco et al. [6] studied a criterium for non-uniqueness which is optimal for the class of binary attractive models.

1.1 Models considered

Although the basic structure of our model is entirely imported from the one of BHS, our notations and terminology differ largely from those of [1].

The process $Z = (Z_t)_{t \in \mathbb{Z}}$ defined in [1] takes values in an alphabet with four symbols, where each symbol is actually a pair, which we denote

$$Z_t = (X_t, \omega_t),$$

with[1] $X_t \in \{+, -\}$, $\omega_t \in \{0, 1\}$. The process can be considered as constructed in two steps. First, a doubly-infinite sequence of i.i.d. random variables $\omega = (\omega_t)_{t \in \mathbb{Z}}$ is sampled, representing the *environment*, with distribution Q:

$$Q(\omega_t = 1) = 1 - Q(\omega_t = 0) = \tfrac{1}{2}.$$

Then, once the environment ω is fixed, a process $X = (X_t)_{t \in \mathbb{Z}}$ is considered, whose conditional distribution given ω is denoted P_ω and called the *quenched distribution*. We will assume that P_ω-almost surely,

$$P_\omega(X_t = \pm \mid X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots) = \tfrac{1}{2}\big\{1 \pm \psi_t^\omega(x_{-\infty}^{t-1})\big\}. \tag{3}$$

where $x_{-\infty}^{t-1} = (x_{t-1}, x_{t-2}, \dots) \in \{\pm\}^{\mathbb{N}}$. The perturbation $\psi_t^\omega : \{\pm\}^{\mathbb{N}} \to [-1, 1]$ describes how the variables of the process X differ from those of an i.i.d. symmetric sequence (which corresponds to $\psi_t^\omega \equiv 0$). The quenched model will always be *attractive*, in the sense that $\psi_t^\omega(x_{-\infty}^{t-1})$ is non-decreasing in each of the variables x_s, $s < t$.

We assume that the functions ψ_t^ω satisfy the following conditions:

C1. For all $x \in \{\pm\}^{\mathbb{N}}$, $\psi_t^\omega(x)$ depends only on the environment variables ω_s, $s \le t$.
C2. The functions ψ_t^ω are odd, $\psi_t^\omega(-x) = -\psi_t^\omega(x)$ for all $x \in \{\pm\}^{\mathbb{N}}$, and bounded uniformly in all their arguments:

$$|\psi_t^\omega(x)| \le \epsilon \quad \text{for some } \epsilon \in (0, 1).$$

C3. The maps $(x, \omega) \mapsto \psi_t^\omega(x)$ are continuous, uniformly in t.
C4. If $\theta : \{0, 1\}^{\mathbb{Z}} \to \{0, 1\}^{\mathbb{Z}}$ denotes the shift, $(\theta\omega)_s := \omega_{s+1}$, then

$$\psi_t^\omega = \psi_0^{\theta_t \omega}.$$

The probability distribution $\mathbb{P}$ of the joint process $Z_t = (X_t, \omega_t)$ is defined as follows. If $A \in \mathcal{F} := \sigma(X_t, t \in \mathbb{Z})$, $B \in \mathcal{G} := \sigma(\omega_t, t \in \mathbb{Z})$, then

$$\mathbb{P}(A \times B) := \int_B P_\omega(A) Q(d\omega). \tag{4}$$

We will sometimes denote $\mathbb{P}$ by $Q \otimes P_\omega$. It can then be verified that under $\mathbb{P}$, $Z = (Z_t)_{t \in \mathbb{Z}}$ is a chain with complete connections specified by the regular g-function

$$g((\pm, \omega_t)|(x_{t-1}, \omega_{t-1}), (x_{t-2}, \omega_{t-2}), \dots) := \tfrac{1}{4}\big\{1 \pm \psi_t^\omega(x_{-\infty}^{t-1})\big\}. \tag{5}$$

Conversely, the distribution of any process Z specified by this g-function can be expressed as in (4). Although the processes specified by g are of a dynamical nature

---

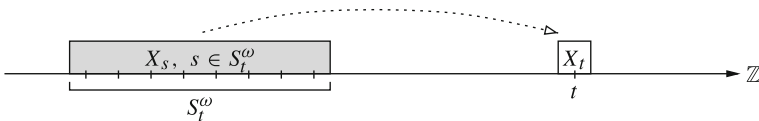[1] Often, we will abbreviate $+1$ (resp. -1) by $+$ (resp. $-$).

Fig. 2 In a given environment ω, the distribution of X_t, conditioned on its past, is determined by the variables X_s, with $s \in S_t^\omega \subset (-\infty, t)$

[the process (x_t, ω_t) at time t having a distribution fixed by the entire past], we will rather be working with the quenched picture in mind, and think only of the variables x_t as being dynamical, evolving in a fixed environment $(\omega_t)_{t \in \mathbb{Z}}$.

The precise definition of the functions ψ_t^ω will be given in Sect. 2.1. Before that we describe, in an informal way, the main ingredients that will appear in their construction.

1.2 Sampling a random set in the past

A natural feature of the model is that the distribution of the quenched process X at time t is determined by its values over a finite (albeit large) region in the past of t. Therefore, for a given environment ω, the starting point will be to associate to each time $t \in \mathbb{Z}$ a random set $S_t = S_t^\omega$ living in the past of t: $S_t \subset (-\infty, t)$. We will say that S_t targets the time t (see Fig. 2). Although each S_t is finite, we will always have, Q-almost surely,

$$\sup_t |S_t| = \infty \quad \text{and} \quad \sup_t \text{dist}(t, S_t) = \infty.$$

In the environment ω, the distribution of X_t conditionned on its past $(X_s)_{s<t}$ [see (3)] is determined by the values of X on S_t. As a matter of fact, the distribution of X_t will depend on the *average* of X on the set S_t:

$$\psi_t^\omega(x_{-\infty}^{t-1}) = \text{ odd function of } \left(\frac{1}{|S_t^\omega|} \sum_{s \in S_t^\omega} x_s \right).$$

The precise dependence will be fixed by some *majority rule*.

Remark 1 In general, S_t^ω will not be an interval; as will be seen below, S_t^ω is defined by a multiscale description of ω, and is made of *unions* of far apart intervals. Nevertheless, we will simplify the figures by picturing S_t as if it were an interval.

The sets S_t will be constructed in such a way that the event pictured on Fig. 3 occurs with positive Q-probability. That event represents a global connectivity satisfied by the sets $(S_t)_{t \in \mathbb{Z}}$ in relation to the origin: 0 is targeted by the set S_0, which we temporarily denote by $S_0(1)$. In turn, all points $s \in S_0(1)$ happen to be targeted by the *same* set, denoted $S_0(2)$. Then, all points $s' \in S_0(2)$ are targeted by the *same* set $S_0(3)$, etc. In this way, for each $j \geq 0$ the variables $\{X_s, s \in S_0(j)\}$, when conditionned on the values of the process on the past of $S_0(j)$, are independent, with a distribution fixed solely by the magnetization of X on $S_0(j + 1)$. In this way, the properties of X in
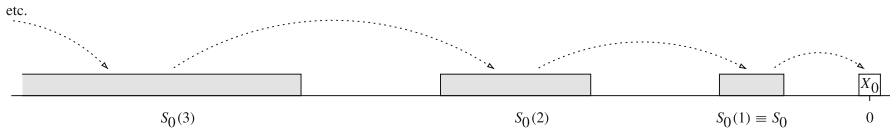
etc.

$$S_0(3) \qquad\qquad\qquad S_0(2) \qquad\qquad S_0(1) \equiv S_0 \qquad 0$$

Fig. 3 An environment in which information is likely to travel from the remote past up to 0

a finite region of $\mathbb{Z}$ will be obtained via values of X on a sequence of sets $S_0(j)$, $j = 1, 2, \ldots$. This sequence will happen to be *multiscale* in the sense that $S_0(j + 1)$ will be orders of magnitude larger than $S_0(j)$. Part of the mechanism will be to obtain estimates on the sizes of these sets. (Obs: The notations of this paragraph will not be used later. For a precise description of the picture just described, see the definition of the event $\{\infty \to k\}$ in Sect. 4.1).

Throughout the paper, $|A|$ will denote the number of elements of A. For simplicity, intervals of $\mathbb{Z}$ will be denoted as $\{a, a + 1, \ldots, b - 1, b\} \equiv [a, b]$. The diameter of $[a, b]$ is $\mathrm{d}([a, b]) := b - a + 1$.

2 The BHS model

The construction of the random sets S_t starts by using the environment ω to partition $\mathbb{Z}$ into blocks of increasing scales. Below, most objects are random and depend on ω, although this will not always be indicated in the notations.

We start by fixing two numbers:

$$\epsilon_* \in (0, 1), \text{ and } k_* \in \mathbb{N}.$$

Later, k_* (the smallest scale) will be chosen large. For all $k \geq k_*$, define

$$\ell_k := \lceil (1 + \epsilon_*)^k \rceil,$$

and let I_k be the word defined as the concatenation of $\ell_k - 1$ symbols "1" followed by a symbol "0":

$$I_k = (1, 1, \ldots, 1, 1, 0). \tag{6}$$

Let $\omega \in \{0, 1\}^{\mathbb{Z}}$ be an environment and $[a, b] \subset \mathbb{Z}$ an interval of diameter ℓ_k. We say that I_k is seen in ω on $[a, b]$ if

$$(\omega_a, \omega_{a+1}, \ldots, \omega_b) = I_k.$$

In a given environment, I_k is seen on infinitely many disjoint intervals (Q-a.s.). Consider two successive occurrences of I_k in ω. That is, suppose I_k is seen on two disjoint intervals $[a, b]$ and $[a', b']$, but not on any other interval contained in $[a, b']$. Then the interval $[b, b' - 1]$ is called a k-block. By definition, a k-block has diameter at least ℓ_k, the first symbol seen on a k-block is 0, and the last $\ell_k - 1$ symbols are 1s.
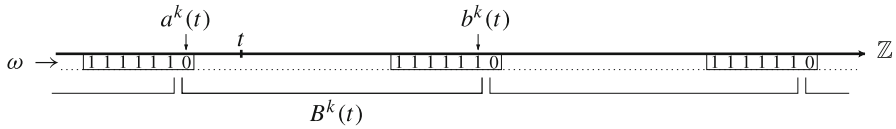
Fig. 4 A partition of Z into k-blocks, using successive occurences of I_k in ω

A given k allows to partition $\mathbb{Z}$ into k-blocks (see Fig. 4): for each $t \in \mathbb{Z}$, there exists a unique k-block containing t, denoted by $B^k(t) = [a^k(t), b^k(t)]$, where $a^k(t)$ [resp. $b^k(t)$] is the leftmost (resp. rightmost) point of $B^k(t)$.

The diameter of a typical k-block is of order (see Lemma 1)

$$\beta_k := 2^{\ell_k}.$$

In a fixed environment, the partition of $\mathbb{Z}$ in k-blocks is coarser than the partition in $(k-1)$-blocks: when $k > k_*$, each k-block B is a disjoint union of one or more $(k-1)$-blocks. If we denote the number of $(k-1)$-blocks in B by $N(B)$, then

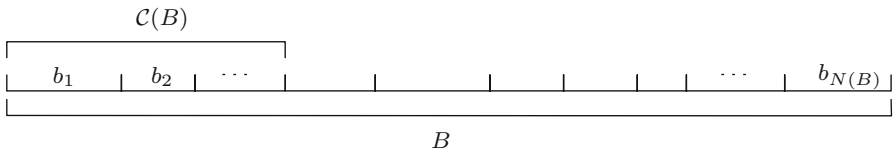$$B = b_1 \cup b_2 \cup \cdots \cup b_{N(B)} \equiv \bigcup_{i=1}^{N(B)} b_i, \qquad (7)$$

where b_1 [resp. $b_{N(B)}$] is the leftmost (resp. rightmost) $(k-1)$-block contained in B. We will verify in Lemma 2 that $N(B)$ is of order

$$\nu_k := \frac{\beta_k}{\beta_{k-1}}.$$

When $k > k_*$, the beginning of a k-block B, decomposed as in (7), is defined as

$$\mathcal{C}(B) := \bigcup_{i=1}^{N(B) \wedge \lfloor \nu_k^{1-\epsilon_*} \rfloor} b_i. \qquad (8)$$

Due to the exponent "$1 - \epsilon_*$" in (8), the beginning of a k-block, when $k > k_*$ is large, is typically smaller than the block itself (see Lemma 3):

For a k_*-block $B = [a, b]$, the beginning is defined in a different manner:

$$\mathcal{C}(B) := \left\{ s \in B : |s - a| \leq \beta_{k_*}^{1+\epsilon_*} \right\}.$$

Since the typical size of a k_*-block B is β_{k_*}, we will verify later that $B = \mathcal{C}(B)$ with high Q-probability.

2.1 The definition of S_t^ω and ψ_t^ω

In order to help understand the precise definition of S_t given below, we first give a definition which is natural but not yet sufficient for our needs.

Fix $t \in \mathbb{Z}$, and consider the first scale for which t is not in the beginning of its block: $k_t := \inf\{k \geq k_* : t \notin \mathcal{C}(B^k(t))\}$. Then, a natural way of defining S_t could be $S_t := \mathcal{C}(B^{k_t}(t))$. Unfortunately, this definition does not guarantee that some event like the one described on Fig. 3 occurs with positive probability. Namely, two distinct points $t', t'' \in S_t$ can very well be targeted by different sets $S_{t'} \neq S_{t''}$. The definitions of S_t and k_t thus need to be modified in some subtle way.

Definition 1 Let $k \geq k_*$. We say that $t \in \mathbb{Z}$ is k-active in the environment ω if for all $j \in \{k_*, \ldots, k\}$,

(1) $t \in \mathcal{C}(B^j(t))$, where $B^j(t) = [a^j(t), b^j(t)]$ is the j-block containing t, and if
(2) $|t - a^j(t)| < \beta_{j+1}$.

Let also $\mathcal{A}_k := \{t \in \mathbb{Z} : t \text{ is } k\text{-active}\}$.

Observe that

$$\mathcal{A}_{k_*} \supset \mathcal{A}_{k_*+1} \supset \cdots \supset \mathcal{A}_k \supset \mathcal{A}_{k+1} \supset \cdots$$

We will see after Lemma 3 that $\mathcal{A}_k \searrow \varnothing$ as $k \to \infty$, Q-almost surely. Therefore, it is natural to define, for $t \in \mathbb{Z}$,

$$k_t = k_t^\omega := \inf\{k \geq k_*; t \notin \mathcal{A}_k\}, \tag{9}$$

with the convention: $\inf \varnothing = \infty$. The set of k-active points inside a k-block B is

$$\mathcal{A}(B) := \mathcal{A}_k \cap B.$$

By definition, $\mathcal{A}(B) \subset \mathcal{C}(B)$. Then, let

$$S_t = S_t^\omega := \begin{cases} \mathcal{A}(B^{k_t}(t)) & \text{if } k_* < k_t < \infty, \\ \varnothing & \text{otherwise.} \end{cases}$$

Unlike in [1], we don't impose S_t to contain an odd number of points. By definition, $S_t \subset (-\infty, t)$, and the two following crucial properties hold:

P1. If $t', t'' \in S_t$, then $S_{t'} = S_{t''}$.
P2. If ω, ω' are such that $\omega_s = \omega'_s$ for all $s \in (\infty, t]$, then $k_t^\omega = k_t^{\omega'}$ and $S_t^\omega = S_t^{\omega'}$.

We can now define ψ_t^ω.

Definition 2 Let $\varphi : [-1, 1] \to [-1, 1]$ be non-decreasing and odd, $\varphi(-z) = -\varphi(z)$, and $h_k > 0$ be a decreasing sequence such that $h_k \searrow 0$ when $k \to \infty$. If $S_t \neq \varnothing$ and $|t - a^{k_t}(t)| < \beta_{k_t+1}$, define

$$\psi_t^\omega(x) := h_{k_t} \varphi\left(\frac{1}{|S_t|} \sum_{s \in S_t} x_s\right). \tag{10}$$

Otherwise,

$$\psi_t^\omega(x) := 0. \tag{11}$$

We check that ψ_t^ω satisfies the properties C1–C4 described earlier. If h_{k_*} is small enough, say $h_{k_*} \leq \frac{1}{2}$, then ψ_t^ω satisfies C2. C3 is guaranteed by the fact that $h_k \searrow 0$ and that a cutoff was introduced so that $\psi_t^\omega = 0$ if $|t - a^{k_t}| \geq \beta_{k_t} + 1$. Then, C4 is clearly satisfied, and C1 is consequence of P2.

We will now present some results concerning the processes Z specified by the g-function defined in (5), with ψ_t^ω defined above. Our interest will be in observing the role played by the behavior of φ at the origin.

2.2 A sharper result for the pure majority rule

In [1], the function φ used is a pure majority rule (see Fig. 1). That is,

$$\varphi_{PMR}(z) := \begin{cases} +1 & \text{if } z \in (0, +1], \\ 0 & \text{if } z = 0, \\ -1 & \text{if } z \in [-1, 0). \end{cases} \tag{12}$$

The behavior of the model then depends crucially on the choice of the sequence h_k. As will be seen, the criterion is roughly the following:

$$\sum_k e^{-h_{k+1}^2 \beta_k^{1-\epsilon_*}} \begin{cases} < \infty & \Rightarrow \text{non-uniqueness}, \\ = \infty & \Rightarrow \text{uniqueness}. \end{cases} \tag{13}$$

The sequence h_k considered in [1] was therefore of the form

$$h_k := \frac{1}{\beta_{k-1}^\alpha}, \quad \alpha > 0. \tag{14}$$

With this particular choice, our first result completes the description given in [1]:

Theorem 1 *Consider the g-function* (5), *with* ψ_t^ω *of the form* (10), *with* φ *discontinuous at the origin like* φ_{PMR}, *and* h_k *defined as in* (14).

(1) *If* $\alpha < \frac{1-\epsilon_*}{2}$, *then there exist two distinct stationary processes* $Z^+ \neq Z^-$ *specified by g.*
(2) *If* $\alpha > \frac{1-\epsilon_*}{2}$, *then there exists a unique stationary process specified by g.*

Item (1) was the main result of [1] but there, the uniqueness regime was not studied. We will actually see (in the proof of Lemma 6) that when $\alpha > \frac{1-\epsilon_*}{2}$, uniqueness holds independently of φ being continuous or not at the origin. The methods presented below don't allow to treat the critical case $\alpha = \frac{1-\epsilon_*}{2}$.

Remark 2 It can be shown (see [1]) that with h_k as in (14), the variation of g satisfies $\text{var}_j(g) \leq c \cdot j^{-\alpha/(1+\epsilon_*)^3}$. Therefore, the Johansson–Öberg criterion (2) guarantees uniqueness when $\alpha > (1 + \epsilon_*)^3/2$. Our result extends uniqueness also to values $\alpha \in \left(\frac{1-\epsilon_*}{2}, \frac{(1+\epsilon_*)^3}{2}\right]$.

2.3 Uniqueness for continuous majority rules

The following two results show, roughly, that any attempt to turn φ continuous at the origin leads to uniqueness.

Theorem 2 *Consider the g-function* (5)*, with ψ_t^ω of the form* (10)*, and assume there exists $\gamma > 0$ such that*

$$\limsup_{z \to 0^+} \frac{\varphi(z)}{z^\gamma} < \infty. \tag{15}$$

If h_k is as in (14)*, then for all $\alpha > 0$ the stationary process specified by g is unique.*

Condition (15) is of course satisfied when φ is differentiable at 0: $\varphi'(0) < \infty$. Some examples of non-uniqueness with $\varphi'(0) = \infty$ for the Bramson-Kalikow model were given in [5]. We will see in Remark 6 that there exist majority rules φ continuous at the origin but which don't satisfy (15), and for which non-uniqueness holds.

Under a stronger condition on φ, we can show that uniqueness holds for all sequence $h_k \searrow 0$.

Theorem 3 *Consider the g-function* (5)*, with ψ_t^ω of the form* (10)*. Assume that φ is locally Lipschitz in a neighborhood of the origin: there exists $\delta > 0$ and $\lambda > 0$ such that*

$$|\varphi(z_2) - \varphi(z_1)| \leq \lambda|z_2 - z_1|, \quad \forall z_1, z_2 \in [-\delta, \delta].$$

Then for any sequence $h_k \searrow 0$, the stationary process specified by g is unique.

An example that leads to uniqueness for all sequence $h_k \searrow 0$ is when $\varphi = \varphi_{\text{lin}}$ is linear at the origin: there exists $0 < \lambda < \infty$ and $\delta > 0$ such that $\varphi_{\text{lin}}(z) = \lambda z$ for all $z \in [-\delta, \delta]$. (This particular example will actually play an important role in the proof.) Otherwise, natural candidates such as $\varphi(z) := \tanh(\beta z)$ also lead to uniqueness even for large $\beta > 0$.
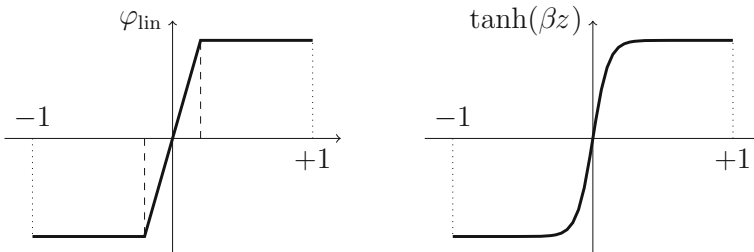
We emphasize that our uniqueness results can't be derived from the classical criteria found in the litterature [such as (2)]. The reason for this is that most criteria are insensitive to the behavior of φ at the origin. In particular, our results allow to build g-functions with arbitrarily slow-decaying variation, that specify a unique stationary process.

The paper is organized as follows. We will first give a detailed description of the environment in Sect. 3, whose spirit follows closely [1]. Nevertheless, our presentation includes a few differences; in particular, our Definition 3 differs from the one in [1]. For that reason, and for the sake of completeness, we will give full proofs. After that, we will describe when an environment should be considered as *good*, and at the beginning of Sect. 4 introduce an event $\{\infty \to k\}$, that will be used constantly in the sequel. We then prove Theorems 1 and 2, using two propositions that are proved later in Sect. 4.6. Theorem 3 is proved in Sect. 5.

3 The environment and properties of blocks

In this section, we study typical properties of a k-block: its diameter, its beginning, and finally we estimate the number of k-active points it contains.

Given a k-block $B = [a, b]$ in some environment ω, we define $\Pi(B)$, the word of B, as the sequence of symbols of ω seen in B:

$$\Pi(B) := (\omega_a, \omega_{a+1}, \ldots, \omega_b).$$

Remember (check Fig. 4) that $\omega_a = 0$, and that $\omega_{b-\ell_k-2} = \cdots = \omega_b = 1$.

To study properties of blocks that only depend on their word, and since the environment is stationary with respect to Q, it will be enough to consider the blocks containing the origin, $B^k(0)$, $k \geq k_*$. The study of $\Pi(B^k(0))$ will be simplified by first studying the block $B^k(0)$ when its first point is fixed at the origin.

Let therefore $\eta = (\eta_i)_{i \geq 1}$ be an i.i.d. sequence, such that $P(\eta_i = 0) = 1 - P(\eta_i = 1) = \frac{1}{2}$. For each $k \geq 1$, we consider the time of first occurrence of I_k [remember (6)] in η, defined by

$$T^k := \inf\big\{j \geq \ell_k : (\eta_{j-\ell_k+1}, \ldots, \eta_j) = I_k\big\}.$$

Defining $\eta_0 := 0$, the random word

$$\Pi^k := (\eta_0, \eta_1, \ldots, \eta_{T^k-1})$$

has the same distribution as of that of a word of a k-block. We call Π^k a k-word, and denote the set of all k-words by $\mathcal{W}^k$. Many notions introduced for k-blocks extend naturally to the k-word Π^k. For instance, the diameter of Π^k, that is the number of symbols it contains, is

$$d(\Pi^k) = T^k. \tag{16}$$

We will study a few elementary properties of the k-word Π^k, and then extend them to $B^k(0)$.

3.1 The diameter

A classical martingale argument (see for instance "the monkey typing Shakespeare" in [11]) allows to compute the expectation of the size of a k-word:

$$E[T^k] = 2^1 + 2^2 + \cdots + 2^{\ell_k}.$$

Therefore,

$$\beta_k \leq E[\mathrm{d}(\Pi^k)] \leq 2\beta_k. \tag{17}$$

We leave it as an exercise to verify that the distribution of $\frac{T^k}{E[T^k]}$ has an exponential tail:

Lemma 1 *For all $j \geq 1$*

$$P(T^k \geq j\beta_k) \leq e^{-j}.$$

Corollary 1 *For all $j \geq 1$,*

$$Q(\mathrm{d}(B^k(0)) \geq j\beta_k) \leq 6je^{-j}. \tag{18}$$

Proof Write $B^k(0) = [a^k(0), b^k(0)]$. For all finite interval $J \subset \mathbb{N}$,

$$Q(\mathrm{d}(B^k(0)) \in J) = \sum_{\substack{\pi \in \mathcal{W}^k \\ \mathrm{d}(\pi) \in J}} Q(\Pi(B^k(0)) = \pi)$$

$$= \sum_{\substack{\pi \in \mathcal{W}^k \\ \mathrm{d}(\pi) \in J}} \sum_{-\mathrm{d}(\pi) < a \leq 0} Q(\Pi(B^k(0)) = \pi, a^k(0) = a). \tag{19}$$

But $\{\Pi(B^k(0)) = \pi, a^k(0) = a\}$, when not empty, is uniquely determined by the following three conditions:

(1) $\omega_j = 1$ for all $j \in \{a - \ell_k + 1, \ldots, a - 1\}$,
(2) $(\omega_a, \omega_{a+1}, \ldots, \omega_{a+\mathrm{d}(\pi)-1}) = \pi$,
(3) $\omega_{a+\mathrm{d}(\pi)} = 0$.

Therefore, by the independence of the variables ω_i,

$$Q(\Pi(B^k(0)) = \pi, a^k(0) = a) = \left(\tfrac{1}{2}\right)^{\ell_k-1} \left(\tfrac{1}{2}\right)^{\mathrm{d}(\pi)} \left(\tfrac{1}{2}\right), \tag{20}$$

which implies

$$Q(\mathrm{d}(B^k(0)) \in J) = \left(\tfrac{1}{2}\right)^{\ell_k} \sum_{\pi \in \mathcal{W}^k \, \mathrm{d}(\pi) \in J} \mathrm{d}(\pi) \left(\tfrac{1}{2}\right)^{\mathrm{d}(\pi)}$$

$$\leq \left(\tfrac{1}{2}\right)^{\ell_k} (\max_{j \in J} j) \sum_{\pi \in \mathcal{W}^k \, \mathrm{d}(\pi) \in J} \left(\tfrac{1}{2}\right)^{\mathrm{d}(\pi)}.$$

But for each $\pi \in \mathcal{W}^k$,

$$P(\Pi^k = \pi) = \left(\tfrac{1}{2}\right)^{d(\pi)-1} \cdot \left(\tfrac{1}{2}\right) = \left(\tfrac{1}{2}\right)^{d(\pi)}, \tag{21}$$

where "$\cdot \left(\tfrac{1}{2}\right)$" appears in order to have a "0" after π, to guarantee the occurrence of the event $\{\Pi^k = \pi\}$. Therefore,

$$\sum_{\substack{\pi \in \mathcal{W}^k \\ d(\pi) \in J}} \left(\tfrac{1}{2}\right)^{d(\pi)} \equiv P\big(d(\Pi^k) \in J\big) = P\big(T^k \in J\big), \tag{22}$$

and we have shown that

$$Q\big(d(B^k(0)) \in J\big) \leq \left(\tfrac{1}{2}\right)^{\ell_k} (\max J) P\big(T^k \in J\big). \tag{23}$$

Let $J_i^k := [i\beta_k, (i+1)\beta_k)$. Since $\left(\tfrac{1}{2}\right)^{\ell_k} \beta_k = 1$,

$$\begin{aligned}
Q\big(d(B^k(0)) \geq j\beta_k\big) &= \sum_{i \geq j} Q\big(d(B^k(0)) \in J_i^k\big) \\
&\leq \sum_{i \geq j}(i+1) P\big(T^k \in J_i^k\big) \leq \sum_{i \geq j}(i+1) P\big(T^k \geq i\beta_k\big).
\end{aligned}$$

Then, (18) follows from Lemma 1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.2 The beginning

When $k > k_*$, Π^k can always be viewed as a concatenation of $(k-1)$-words:

$$\Pi^k = \Pi_1^{k-1} \star \cdots \star \Pi_{N(\Pi^k)}^{k-1},$$

where $\Pi_1^{k-1} := \Pi^{k-1}$ and

$$\Pi_j^{k-1} := \Pi^{k-1} \circ \theta_{d(\Pi_1^{k-1} \star \cdots \star \Pi_{j-1}^{k-1})}.$$

The number of $(k-1)$-words contained in Π^k, $N(\Pi^k)$, is geometric:

Lemma 2 *For all $k > k_*$, there exists $p_k \in (0, 1)$, $v_k^{-1} \leq p_k \leq 2v_k^{-1}$, such that*

$$\forall j \geq 1, \quad P(N(\Pi^k) = j) = (1 - p_k)^{j-1} p_k. \tag{24}$$

In particular, $E[N(\Pi^k)] = p_k^{-1}$.

Proof Consider an independent identically distributed sequence of $(k-1)$-words, with the same distribution as Π^{k-1}: $\Pi_1^{k-1}, \Pi_2^{k-1}, \ldots$. When sampling a $(k-1)$-word, we say that this $(k-1)$-word is closing if the first occurrence of I_{k-1} coincides with the first occurrence of I_k, which means that the first occurrence of I_{k-1} is preceded by a sequence of $\ell_k - \ell_{k-1}$ symbols "1". Therefore, the concatenation of the first j $(k-1)$-words of the sequence $\Pi_1^{k-1}, \Pi_2^{k-1}, \ldots$ is a k-word if and only if the $(j-1)$-th first are not closing and the j-th is closing. Defining

$$p_k := P(\Pi^{k-1} \text{ is closing}) = P(T^{k-1} = T^k),$$

we obtain (24). If $T^{k-1} = t \geq \ell_k$, we denote by r_t the word seen in the interval $[t - \ell_k + 1, t - \ell_{k-1}]$, of diameter $\ell_k - \ell_{k-1}$, and by q_t the word seen in the interval $[t - \ell_{k-1} + 1, t]$. We have

$$
\begin{aligned}
P(T^{k-1} = T^k = t) &= P(r_t = (1, \ldots, 1), T^{k-1} = t) \\
&= P(T^{k-1} > t - \ell_k, r_t = (1, \ldots, 1), q_t = I_{k-1}) \\
&= P(T^{k-1} > t - \ell_k) P(r_t = (1, \ldots, 1)) P(q_t = I_{k-1}) \\
&= \beta_k^{-1} P(T^{k-1} > t - \ell_k).
\end{aligned}
$$

Therefore,

$$p_k = \sum_{t \geq \ell_k} P(T^{k-1} = T^k = t) = \beta_k^{-1} E[T^{k-1}].$$

Using (17), we get $v_k^{-1} \leq p_k \leq 2 v_k^{-1}$. $\square$

All notions previously defined for blocks, such as active points, "being good", etc, have immediate analogs for words. Namely, any k-word $\Pi \in \mathcal{P}^k$, can be identified as $\Pi(B^k(0))$, where $B^k(0)$ is assumed to have its first point pinned at the origin: $a^k(0) = 0$. The beginning of Π^k is the interval

$$\mathcal{C}(\Pi^k) := \left[0, \sum_{i=1}^{N(\Pi^k) \wedge \lfloor v_k^{1-\epsilon_*} \rfloor} d(\Pi_i^{k-1}) \right).$$

Using (17),

$$E[d(\mathcal{C}(\Pi^k))] \leq v_k^{1-\epsilon_*} E[d(\Pi^{k-1})] \leq 2 v_k^{1-\epsilon_*} \beta_{k-1}. \tag{25}$$

We can now study the position of a point relative to the beginning of each of the k-blocks in which it is contained:

Lemma 3 *Let $t \in \mathbb{Z}$. For $k = k_*$,*

$$Q\big(t \in \mathcal{C}(B^{k_*}(t))\big) \geq 1 - 6\lfloor \beta_{k_*}^{\epsilon_*} \rfloor e^{-\lfloor \beta_{k_*}^{\epsilon_*} \rfloor}. \tag{26}$$

For all $k > k_*$,

$$Q\big(t \in \mathcal{C}(B^k(t))\big) \leq 2\nu_k^{-\epsilon_*}, \tag{27}$$

Remark 3 Observe that ν_k diverges superexponentially in k, and so (27) implies

$$\sum_{k > k_*} Q(t \in \mathcal{C}(B^k(t)) < \infty.$$

Therefore, by Borel–Cantelli's Lemma, $t \notin \mathcal{C}(B^k(t))$ for all large enough k. As a consequence, $Q(k_t < \infty) = 1$.

Proof of Lemma 3 It suffices to consider $t = 0$. On the one hand, by Corollary 1,

$$\begin{aligned}
Q\big(0 \notin \mathcal{C}(B^{k_*}(0))\big) &\leq Q\big(B^{k_*}(0)\backslash\mathcal{C}(B^{k_*}(0)) \neq \varnothing\big) \\
&\leq Q\big(\mathrm{d}(B^k(0)) \geq \beta_{k_*}^{1+\epsilon_*}\big) \leq 6\lfloor\beta_{k_*}^{\epsilon_*}\rfloor e^{-\lfloor\beta_{k_*}^{\epsilon_*}\rfloor}.
\end{aligned}$$

On the other hand, using (20), (21), (25),

$$\begin{aligned}
Q(0 \in \mathcal{C}(B^k(0))) &= \sum_{a \leq 0} Q(0 \in \mathcal{C}(B^k(0)), a^k(0) = a) \\
&= \sum_{a \leq 0} \sum_{\substack{\pi \in \mathcal{W}^k: \\ \mathrm{d}(\mathcal{C}(\pi)) > |a|}} Q(\Pi(B^k(0)) = \pi, a^k(0) = a) \\
&= \big(\tfrac{1}{2}\big)^{\ell_k} \sum_{\pi \in \mathcal{W}^k} \mathrm{d}(\mathcal{C}(\pi))\big(\tfrac{1}{2}\big)^{\mathrm{d}(\pi)} = \beta_k^{-1} E[\mathrm{d}(\mathcal{C}(\Pi^k))] \leq 2\nu_k^{-\epsilon_*}.
\end{aligned}$$

$\square$

3.3 The number of active points

Definition 3 A k-block B is good if

(1) $\mathrm{d}(B) < \frac{1}{2}\beta_k^{1+\epsilon_*}$, and if
(2) $n_1(k) := \beta_k^{1-\epsilon_*}2^{-k} < |\mathcal{A}(B)| < n_2(k) := \beta_k^{1-\epsilon_*}\beta_{k-1}^{2\epsilon_*}$.

If not good, B is bad.

Proposition 1 *If k_* is large enough, then for all $k > k_*$,*

$$Q(B^k(0) \text{ is bad}) \leq 2^{-k}. \tag{28}$$

Since the event $\{B^k(0) \text{ is good}\}$ is determined by the word of $B^k(0)$, we first obtain a similar result for words.

The notion of "good" extends naturally to k-words. The set of good k-words is denoted $\mathcal{W}_{\text{good}}^k$, and $\mathcal{W}_{\text{bad}}^k := \mathcal{W}^k \backslash \mathcal{W}_{\text{good}}^k$.

Lemma 4 *If k_* if large enough, then for all $k > k_*$,*

$$P\big(\Pi^k \in \mathcal{W}^k_{bad}\big) \leq 3 \cdot 2^{-3k} + 2\beta^{-\epsilon_*}_{k-1}. \tag{29}$$

Proof We write $\mathcal{W}^k_{good} = \mathcal{W}^{k,1}_{good} \cap \mathcal{W}^{k,2}_{good}$, where

$$\mathcal{W}^{k,1}_{good} := \big\{\pi \in \mathcal{W}^k : d(\pi) < \tfrac{1}{2}\beta^{1+\epsilon_*}_k \text{ and } |\mathcal{A}(\pi)| > n_1(k)\big\},$$
$$\mathcal{W}^{k,2}_{good} := \big\{\pi \in \mathcal{W}^k : |\mathcal{A}(\pi)| < n_2(k)\big\}.$$

Let then $\mathcal{W}^{k,i}_{bad} := \mathcal{W}^k \backslash \mathcal{W}^{k,i}_{good}$. We first prove that for all $k \geq k_*$,

$$P\big(\Pi^k \in \mathcal{W}^{k,1}_{bad}\big) \leq 3 \cdot 2^{-3k}. \tag{30}$$

We will proceed by induction on k. Let k_* be large enough, such that for all $k \geq k_*$,

$$e^{-\tfrac{1}{2}\beta^{\epsilon_*}_k} \leq 2^{-3k}, \quad 2v^{-\epsilon_*}_k \leq 2^{-3k}.$$

(Observe that $k_* \nearrow \infty$ as $\epsilon_* \searrow 0$.) We start with the case $k = k_*$:

$$P\big(\Pi^{k_*} \in \mathcal{W}^{k_*,1}_{bad}\big) \leq P\big(d(\Pi^{k_*}) \geq \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}\big)$$
$$+ P\big(d(\Pi^{k_*}) < \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}, |\mathcal{A}(\Pi^{k_*})| \leq n_1(k_*)\big).$$

By Lemma [1],

$$P\big(d(\Pi^{k_*}) \geq \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}\big) = P\big(T^{k_*} \geq \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}\big) \leq e^{-\tfrac{1}{2}\beta^{\epsilon_*}_{k_*}} \leq 2^{-3k_*}. \tag{31}$$

On the other hand, $\mathcal{A}(\Pi^{k_*})$ is an interval, and $d(\Pi^{k_*}) < \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}$ implies that $|\mathcal{A}(\Pi^{k_*})| = d(\Pi^{k_*}) = T^{k_*}$. Therefore,

$$P\big(d(\Pi^{k_*}) < \tfrac{1}{2}\beta^{1+\epsilon_*}_{k_*}, |\mathcal{A}(\Pi^{k_*})| \leq n_1(k_*)\big) \leq P\big(T^{k_*} \leq n_1(k_*)\big)$$
$$\leq n_1(k_*)\big(\tfrac{1}{2}\big)^{\ell_{k_*}} = \beta^{-\epsilon_*}_{k_*} 2^{-k_*} \leq 2 \cdot 2^{-3k_*}.$$

Therefore, (30) is proved for $k = k_*$. Suppose that (30) holds for $k - 1$. Remember that Π^k is a concatenation of $(k-1)$-words, denoted by Π_j, $j = 1, \ldots, N(\Pi^k)$. We define the events:

$$A_1 := \{d(\Pi^k) < \tfrac{1}{2}\beta^{1+\epsilon_*}_k\},$$
$$A_2 := \{N(\Pi^k) > v^{1-\epsilon_*}_k\},$$
$$A_3 := \left\{\begin{array}{c}\text{at least half of the } (k-1)\text{-words} \\ \text{in } \mathcal{C}(\Pi^k) \text{ are in } \mathcal{W}^{k-1,1}_{good}\end{array}\right\}.$$

We claim that $A_1 \cap A_2 \cap A_3 \subset \{\Pi^k \in \mathcal{W}_{\mathrm{good}}^{k,1}\}$. Indeed, A_1 ensures that the first condition in $\mathcal{W}_{\mathrm{good}}^{k,1}$ is satisfied. Furthermore, in $A_1 \cap A_2 \cap A_3$, every $(k-1)$-word $\Pi_j \subset \mathcal{C}(\Pi^k)$ is at distance $\leq \frac{1}{2}\beta_k^{1+\epsilon_*} \leq \beta_{k+1}$ of the origin, and therefore, each active point of Π_j is active in Π^k. As a consequence,

$$
\begin{aligned}
|\mathcal{A}(\Pi^k)| &= \sum_{\Pi_j \subset \mathcal{C}(\Pi^k)} |\mathcal{A}(\Pi_j)| \\
&\geq \sum_{\substack{\Pi_j \subset \mathcal{C}(\Pi^k): \\ \Pi_j \in \mathcal{W}_{\mathrm{good}}^{k-1,1}}} |\mathcal{A}(\Pi_j)| > n_1(k-1) \sum_{\substack{\Pi_j \subset \mathcal{C}(\Pi^k): \\ \Pi_j \in \mathcal{W}_{\mathrm{good}}^{k-1,1}}} 1 \geq n_1(k-1) \cdot \tfrac{1}{2}v_k^{1-\epsilon_*} \equiv n_1(k).
\end{aligned}
$$

Therefore, $\Pi^k \in \mathcal{W}_{\mathrm{good}}^{k,1}$. It follows that

$$
P\big(\Pi^k \in \mathcal{W}_{\mathrm{bad}}^{k,1}\big) \leq P\big(A_1^c\big) + P\big(A_2^c\big) + P\big(A_2 \cap A_3^c\big).
$$

As in (31), $P(A_1^c) \leq 2^{-3k}$. By Lemma 2,

$$
\begin{aligned}
P(A_2^c) &= \sum_{j=1}^{\lfloor v_k^{1-\epsilon_*} \rfloor} P(N(\Pi^k) = j) \\
&= \sum_{j=1}^{\lfloor v_k^{1-\epsilon_*} \rfloor} (1-p_k)^{j-1} p_k \leq v_k^{1-\epsilon_*} p_k \leq 2v_k^{-\epsilon_*} \leq 2^{-3k}.
\end{aligned}
$$

On A_2, $\mathcal{C}(\Pi^k)$ contains exactly $\lfloor v_k^{1-\epsilon_*} \rfloor$ $(k-1)$-words. Therefore, using the induction hypothesis (30) for $k-1$,

$$
\begin{aligned}
P(A_2 \cap A_3^c) &= P\left(A_2 \cap \left\{\begin{array}{c}\text{at least half of the } (k-1)\text{-words} \\ \text{of } \mathcal{C}(\Pi^k) \text{ are in } \mathcal{W}_{\mathrm{bad}}^{k-1,1}\end{array}\right\}\right) \\
&\leq P\left(\left\{\begin{array}{c}\text{at least half of the } \lfloor v_k^{1-\epsilon_*} \rfloor (k-1)\text{-words} \\ \text{of } \mathcal{C}(\Pi^k) \text{ are in } \mathcal{W}_{\mathrm{bad}}^{k-1,1}\end{array}\right\}\right) \\
&\leq \sum_{j=\frac{1}{2}\lfloor v_k^{1-\epsilon_*} \rfloor}^{\lfloor v_k^{1-\epsilon_*} \rfloor} \binom{\lfloor v_k^{1-\epsilon_*} \rfloor}{j} \left(3 \cdot 2^{-3(k-1)}\right)^j \leq \left(2^{-3(k-1)+6}\right)^{\frac{1}{2}\lfloor v_k^{1-\epsilon_*} \rfloor} < 2^{-3k}.
\end{aligned}
$$

This proves (30). It remains to prove that for all $k > k_*$,

$$
P\big(\Pi^k \in \mathcal{W}_{\mathrm{bad}}^{k,2}\big) \leq 2\beta_{k-1}^{-\epsilon_*}. \tag{32}
$$

Since $\mathcal{A}(\Pi^k) \subset \mathcal{C}(\Pi^k)$, we estimate $\mathrm{d}(\mathcal{C}(\Pi^k))$. We define $T_1^{k-1} := T^{k-1}$ and for $i > 1$,

$$
T_i^{k-1} := \inf\big\{j > T_{i-1}^{k-1} : (\eta_{j-\ell_k+1}, \ldots, \eta_j) = I_k\big\}.
$$

The increments $\tau_i^{k-1} := T_i^{k-1} - T_{i-1}^{k-1}$ are i.i.d., and $E[\tau_1^{k-1}] = E[T_1^{k-1}] \leq 2\beta_{k-1}$. Moreover, $d(\mathcal{C}(\Pi^k)) \leq \sum_{i=1}^{\lfloor v_k^{1-\epsilon_*} \rfloor} \tau_i^{k-1}$. Therefore,

$$P\big(|\mathcal{A}(\Pi^k)| \geq n_2(k)\big) \leq P\left(\sum_{i=1}^{\lfloor v_k^{1-\epsilon_*} \rfloor} \tau_i^{k-1} \geq n_2(k)\right) \leq \frac{E[\tau_1^{k-1}]}{\beta_{k-1}^{1+\epsilon_*}} \leq 2\beta_{k-1}^{-\epsilon_*}.$$

This proves (32). Together, (30) and (32) give (29). $\qquad\square$

Proof of Proposition 1 Take $k > k_*$, where k_* was defined in the proof of Lemma 4. We have

$$\begin{aligned}
Q(B^k(0) \text{ is bad}) &\leq Q\big(d(B^k(0)) \geq k\beta_k\big) \\
&\quad + Q\big(B^k(0) \text{ is bad}, d(B^k(0)) \leq k\beta_k\big).
\end{aligned} \tag{33}$$

By Corollary 1, $Q\big(d(B^k(0)) \geq k\beta_k\big) \leq 6ke^{-k}$. Then,

$$\begin{aligned}
Q\big(B^k(0) \text{ is bad}, d(B^k(0)) \leq k\beta_k\big) &= \left(\tfrac{1}{2}\right)^{\ell_k} \sum_{\substack{\pi \in \mathcal{W}_{\text{bad}}^k \\ d(\pi) \leq k\beta_k}} d(\pi) \left(\tfrac{1}{2}\right)^{d(\pi)} \\
&\leq k \sum_{\pi \in \mathcal{W}_{\text{bad}}^k} \left(\tfrac{1}{2}\right)^{d(\pi)} \equiv k P\big(\Pi^k \in \mathcal{W}_{\text{bad}}^k\big).
\end{aligned}$$

Using Lemma 4,

$$Q\big(B^k(0) \text{ is bad}\big) \leq 6ke^{-k} + k(3 \cdot 2^{-3k} + 2\beta_{k-1}^{-\epsilon_*}). \tag{34}$$

Taking k_* large enough, this proves (28). $\qquad\square$

4 Proofs of Theorems 1 and 2

The proofs of all the results will study the process X under the quenched measure P_ω, using environments ω for which the influence of the remote past on the present (for example on a local event like $\{X_0 = +\}$) can be computed and related to φ and to the sequence h_k.

4.1 The event $\{\infty \rightarrow k\}$

To start, consider a set of variables $\{X_s, s \in R\}$, where R is a finite subset of $\mathbb{Z}$. There clearly exists some $k(1) \geq k_*$ such that $R \subset B^{k(1)}(0)$. Furthermore, using Remark 3, we can take $k(1)$ sufficiently large, and guarantee that $R \subset B^{k(1)}(0) \backslash \mathcal{C}(B^{k(1)}(0))$. But then, by the definition of the g-function constructed with ψ_t^ω, the only way by which
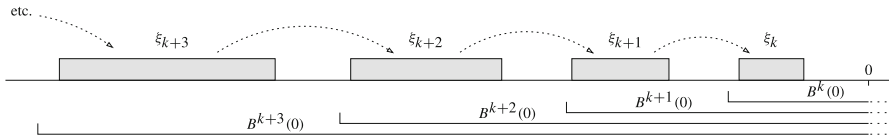
Fig. 5 The quenched distribution of the process X, on the event $\{\infty \to k\}$

the remote past influences the variables in R is through the value of the average of the variables $\{X_t, \ t \in \mathcal{A}(B^{k(1)}(0))\}$.

Repeating the same procedure with $B^{k(1)}(0)$ in place of R, we deduce that the distribution of $\{X_t, \ t \in \mathcal{A}(B^{k(1)}(0))\}$ is entirely determined by the values of $\{X_t, \ t \in \mathcal{A}(B^{k(2)}(0))\}$ for some sufficiently large $k(2)$, etc.

Our aim will be to ensure that the random sequence $k(i)$ satisfies $k(i+1) = k(i)+1$ for all large i, and that the sizes of the sets $\mathcal{A}(B^{k(i)}(0))$ are under control. We thus define, for all $k > k_*$,

$$\{\infty \to k\} := \bigcap_{j \geq k} \left\{0 \notin \mathcal{C}(B^j(0)), \ B^j(0) \text{ is good}\right\}. \tag{35}$$

The notation used suggests that the event is of the type described earlier in Fig. 3. Indeed, let $\omega \in \{\infty \to k\}$. Take $j \geq k$, and $t \in \mathcal{A}(B^j(0))$. Since $0 \notin \mathcal{C}(B^{j+1}(0))$ and $B^{j+1}(0) = B^{j+1}(t)$, we have that $t \notin \mathcal{C}(B^{j+1}(t))$ which implies $t \notin \mathcal{A}(B^{j+1}(t))$. Therefore, $k_t = j + 1$. Moreover, since $B^{j+1}(0)$ is good, we have that $d(B^{j+1}(0)) \leq \frac{1}{2}\beta_{j+1}^{1+\epsilon_*} \leq \beta_{j+2}$. This implies that $S_t = \mathcal{A}(B^{j+1}(0))$, and

$$\psi_t^\omega = h_{j+1}\varphi\left(\frac{1}{|\mathcal{A}(B^{j+1}(0))|} \sum_{s \in \mathcal{A}(B^{j+1}(0))} X_s\right). \tag{36}$$

Therefore, on $\{\infty \to k\}$, for all $j \geq k$, the variables $\{X_t, t \in \mathcal{A}(B^j(0))\}$ are i.i.d., and their distribution is fixed by the value of the magnetization of $\{X_t, t \in \mathcal{A}(B^{j+1}(0))\}$. That is, the distribution of the process X on any finite region is related to the behavior of the non-homogeneous Markov sequence

$$\xi_j := \frac{1}{|\mathcal{A}(B^j(0))|} \sum_{s \in \mathcal{A}(B^j(0))} X_s, \quad j \geq k.$$

The transition probability of the chain will be studied using the following relation, which holds on the event $\{\infty \to k\}$, for all $j \geq k$ (Fig. 5):

$$E_\omega[\xi_j \mid \xi_{j+1}] = h_{j+1}\varphi(\xi_{j+1}), \tag{37}$$

Proposition 2 *Let $k \geq k_*$. There exists $\lambda(\epsilon_*, k) > 0$ with $\lambda(\epsilon_*, k) \searrow 0$ when $k \to \infty$, such that*

$$Q(\infty \to k) \geq 1 - \lambda(\epsilon_*, k). \tag{38}$$

Moreover, there exists a random scale $K = K(\omega)$, $Q(K < \infty) = 1$, such that

$$Q(\infty \to K) = 1.$$

Proof If k_* is as large as in Proposition 1, then for $k \geq k_*$

$$Q(\{\infty \to k\}^c) \leq \sum_{j \geq k} \{Q(B^j(0) \text{ is bad}) + Q(0 \in \mathcal{C}(B^j(0)))\}$$

$$\leq \sum_{j \geq k} \{2^{-j} + 2v_j^{-\epsilon_*}\},\tag{39}$$

which is summable in k. The existence of K follows by Borel-Cantelli's Lemma. $\square$

4.2 The measures P_ω^+, P_ω^- and their maximal coupling

Our proofs will rely on the use of two particular processes specified by g, $Z^+ = (X^+, \omega)$ and $Z^- = (X^-, \omega)$, symmetric with respect to each other in the sense that

$$\mathbb{P}(X_s^- = +) = \mathbb{P}(X_s^+ = -).\tag{40}$$

Each $X^\#$ will actually be the coordinate process associated to a probability measure $P_\omega^\#$ on $\{\pm\}^{\mathbb{Z}}$ constructed with a pure boundary condition $\# \in \{+, -\}$. The construction given below is standard.

For $x, y \in \{\pm\}^{\mathbb{Z}}$, let $x_s^t := (x_t, \ldots, x_s)$ and $x_{-\infty}^t := (x_t, x_{t-1}, \ldots)$. If $\omega \in \{0, 1\}^{\mathbb{Z}}$, define

$$g_t^\omega(x_t | x_{-\infty}^{t-1}) := \tfrac{1}{2}\{1 + x_t \psi_t^\omega(x_{-\infty}^{t-1})\}.$$

Define also the cylinder $[x]_s^t := \{y \in \{\pm\}^{\mathbb{Z}} : y_i = x_i, \ i \in [s, t]\}$ and $x_s^t y_{-\infty}^{s-1} := (x_t, \ldots, x_s, y_{s-1}, y_{s-2}, \ldots)$. For each $N \in \mathbb{N}$, $\eta \in \{\pm\}^{\mathbb{Z}}$, we define a probability measure on $\{\pm\}^{(-N, \infty)}$ by setting

$$P_\omega^{\eta, N}([x]_{-N+1}^s) := g_{-N+1}^\omega(x_{-N+1} | \eta_{-\infty}^{-N}) \prod_{t=-N+2}^s g_t^\omega(x_t | x_{-N+1}^{t-1} \eta_{-\infty}^{-N}).$$

When $\eta^1 \leq \eta^2$ (pointwise), $P_\omega^{\eta^1, N}$ and $P_\omega^{\eta^2, N}$ can be coupled as follows. Consider an i.i.d. sequence of random variables $(U_t)_{t > -N}$, each with a uniform distribution on $[0, 1]$. We construct two processes, X^1 and X^2, through a sequence of pairs, $\Delta_t = \binom{X_t^2}{X_t^1}$, $t > -N$, in such a way that $(X_t^\#)_{t > -N}$ has distribution $P_\omega^{\eta^\#, N}$ and such that $X_t^1 \leq X_t^2$ for all $t > -N$. For $s \leq -N$, set $x_s^\# := \eta_s^\#$. Assume that the pairs $\Delta_s = \binom{x_s^2}{x_s^1}$ have

been sampled for all $s < t$, and that these satisfy $x_s^1 \leq x_s^2$. Let

$$\Delta_t = \begin{pmatrix} X_t^2 \\ X_t^1 \end{pmatrix} := \begin{pmatrix} + \\ - \end{pmatrix} 1_{A_t} + \begin{pmatrix} + \\ + \end{pmatrix} 1_{B_t} + \begin{pmatrix} - \\ + \end{pmatrix} 1_{C_t}, \tag{41}$$

where

$$\begin{aligned}
A_t &:= \left\{ 0 \leq U_t < g_t^\omega(+ \mid (x^2)_{-\infty}^{t-1}) - g_t^\omega(+ \mid (x^1)_{-\infty}^{t-1}) \right\}, \\
B_t &:= \left\{ g_t^\omega(+ \mid (x^2)_{-\infty}^{t-1}) - g_t^\omega(+ \mid (x^1)_{-\infty}^{t-1}) \leq U_t < g_t^\omega(+ \mid (x^2)_{-\infty}^{t-1}) \right\}, \\
C_t &:= \left\{ g_t^\omega(+ \mid (x^2)_{-\infty}^{t-1}) \leq U_t \leq 1 \right\}.
\end{aligned} \tag{42}$$

We of course have $\mathsf{P}(X_t^1 \leq X_t^2) = 1$, and

$$\begin{aligned}
\mathsf{P}\big(X_t^2 = + \mid X_{t-1}^2 = x_{t-1}^2, \ldots, X_{-N+1}^2 = x_{-N+1}^2\big) &= \mathsf{P}\big(A_t \cup B_t\big) \\
&= g_t^\omega\big(+ \mid (x^2)_{-\infty}^{t-1}\big),
\end{aligned}$$

and so the distribution of $(X_t^2)_{t>-N}$ is given by $P_\omega^{\eta^2, N}$. Similarly,

$$\begin{aligned}
\mathsf{P}\big(X_t^1 = + \mid X_{t-1}^1 = x_{t-1}^1, \ldots, X_{-N+1}^1 = x_{-N+1}^1\big) &= \mathsf{P}\big(B_t\big) \\
&= g_t^\omega\big(+ \mid (x^1)_{-\infty}^{t-1}\big),
\end{aligned}$$

and so the distribution of $(X_t^1)_{t>-N}$ is given by $P_\omega^{\eta^1, N}$.

The above coupling allows to extract information about the measures $P_\omega^{\eta^\#, N}$. First, $x^1 \leq x^2$ implies $g_t^\omega(+ \mid (x^1)_{-\infty}^{t-1}) \leq g_t^\omega(+ \mid (x^2)_{-\infty}^{t-1})$, and so we always have

$$P_\omega^{\eta^1, N}(X_t = +) \leq P_\omega^{\eta^2, N}(X_t = +), \quad t > -N.$$

More generally, if $f : \{\pm\}^{(-N+1, \infty)} \to \mathbb{R}$ is an increasing local function (that is: non-decreasing in each variable x_s and depending only on a finite number of coordinates), then

$$E_\omega^{\eta^1, N}[f] \leq E_\omega^{\eta^2, N}[f]. \tag{43}$$

Using the previous item, one can also construct two processes P_ω^+ and P_ω^- by taking monotone limits. Namely, let $P_\omega^{+, N}$ be constructed as above using the boundary condition $\eta_s \equiv +$ for all s. Using (43), it is easy to see that for all local increasing f,

$$E_\omega^{+,N+1}[f] \leq E_\omega^{+,N}[f].$$

which allows to define $E_\omega^+[f] := \lim_{N\to\infty} E_\omega^{+,N}[f]$. Since this extends to all continuous function, it defines a measure P_ω^+. It can then be verified that the coordinate process $X = (X_t)_{t\in\mathbb{Z}}$ defined by $X_t(x) := x_t$ satisfies (3) (with P_ω^+ in place of P_ω).

4.3 Non-uniqueness when $\alpha < \frac{1-\epsilon_*}{2}$

Let $\xi_k^\#$ denote the average of $X^\#$ over $\mathcal{A}(B^k(0))$ (if $\mathcal{A}(B^k(0)) = \varnothing$, let $\xi_k^\# := 0$). To obtain non-uniqueness, we will show that when $\alpha < \frac{1-\epsilon_*}{2}$,

$$\mathbb{P}(\xi_k^+ > 0) > \tfrac{1}{2} > \mathbb{P}(\xi_k^- > 0), \tag{44}$$

for all large enough k. Actually, due to the attractiveness of g, the following lower bound holds for all ω:

$$P_\omega^+(\xi_k \geq 0) \geq \tfrac{1}{2}. \tag{45}$$

Proposition 3 *Under the hypotheses of Theorem 1, with $\alpha < \frac{1-\epsilon_*}{2}$, there exists for all $k' > k_*$ some $\epsilon(k')$, $\epsilon(k') \searrow 0$ as $k' \to \infty$, such that*

$$\forall\, \omega \in \{\infty \to k'\}, \quad P_\omega^+(\xi_{k'} > 0) \geq 1 - \epsilon(k'). \tag{46}$$

Proof of Theorem 1, item (1) Using (46),

$$\mathbb{P}(\xi_k^+ > 0) \geq \int_{\{\infty \to k\}} P_\omega^+(\xi_k > 0) Q(d\omega)$$
$$\geq (1 - \epsilon(k)) Q(\infty \to k).$$

By taking k large, this lower bound is $> \frac{1}{2}$. This proves (44), and thereby item (1) of Theorem 1.   □

Remark 4 As the proof of Proposition 3 will show, it is possible to distinguish X^+ and X^- even at the origin. Namely, it can be shown that

$$\forall\, \omega \in \{\infty \to k_* + 1\}, \quad P_\omega^+(X_0 = +) \geq \tfrac{1}{2} + \tau, \tag{47}$$

where $\tau > 0$ once k_* is taken large depending on α. Then, by (45) and (47),

$$\mathbb{P}(X_0^+ = +) \geq \left(\tfrac{1}{2} + \tau\right) Q(\infty \to k_* + 1) + \tfrac{1}{2} Q(\{\infty \to k_* + 1\}^c)$$
$$= \tfrac{1}{2} + \tau Q(\infty \to k_* + 1)$$
$$> \tfrac{1}{2}.$$

Nevertheless, we prefer to avoid having k_* depend on α.

Remark 5 In [1], non-uniqueness was obtained by showing that when $\alpha < \frac{1-\epsilon_*}{2}$, any process $\mathbb{P}$ specified by g must satisfy

$$\mathbb{P}\left(\left\{\lim_{k\to\infty} 1_{\{\xi_k>0\}} = 1\right\} \cup \left\{\lim_{k\to\infty} 1_{\{\xi_k<0\}} = 1\right\}\right) = 1.$$

From this, the existence of two distinct processes can be deduced, using an argument based on symmetry and ergodic decomposition.

4.4 The signature of uniqueness

Assume the environment ω is typical, in the sense that $\{\infty \to k'\}$ occurs for some large enough k'. As we have seen, the distribution of X on any finite region can be studied via the information contained in the sequence ξ_k, $k \geq k'$. On the one hand, we have seen in (44) that non-uniqueness is observed through some asymmetry in the distribution of ξ_k when k is large. Uniqueness, on the other hand, will essentially be characterized by showing that the variables ξ_k are symmetric:

$$E_\omega[\xi_k] = 0 \quad \text{for all large } k. \tag{48}$$

Observe that regardless of the details of φ,

$$\lim_{k\to\infty} E_\omega[\xi_k] = 0 \tag{49}$$

always holds. Namely, for all $k \geq k'$,

$$E_\omega[\xi_k] = E_\omega\big[E_\omega[\xi_k \mid \xi_{k+1}]\big] = h_{k+1} E_\omega[\varphi(\xi_{k+1})] = O\big(h_{k+1}\big).$$

More can be said: when conditioned on ξ_{k+1}, ξ_k is a Bernoulli sum of i.i.d. variables X_s with expectation $h_{k+1}\varphi(\xi_{k+1})$. Therefore, for any fixed $\epsilon > 0$, if k is large enough so that $h_{k+1} \leq \epsilon/2$, a standard large deviation estimate yields

$$P_\omega\big(|\xi_k| > \epsilon \mid \xi_{k+1}\big) \leq e^{-c|\mathcal{A}(B^k(0))|} \leq e^{-cn_1(k)}, \tag{50}$$

where $c = c(\epsilon) > 0$ (we have used the fact that $B^k(0)$ is good). Therefore,

$$\forall \epsilon > 0, \quad P_\omega\big(|\xi_k| \leq \epsilon \text{ for all large enough } k\big) = 1. \tag{51}$$

Since the variables ξ_k almost surely tend to zero when $k \to \infty$, observing some (a)symmetry in their distribution is a delicate problem.

Uniqueness will be obtained with the help of the following criterion, whose proof can be found in Appendix A.

Theorem 4 *Assume that*

$$E_\omega^+[X_t] = 0 = E_\omega^-[X_t] \quad \forall\, t \in \mathbb{Z}. \tag{52}$$

Then, $P_\omega^+ = P_\omega^-$, and any measure P_ω satisfying (3) coincides with P_ω^+ and P_ω^-.

4.5 Uniqueness when $\alpha > \frac{1-\epsilon_*}{2}$

Proposition 4 *Let $\mathbb{P} = Q \otimes P_\omega$ be the distribution of any process specified by g. Under the hypotheses of Theorem 1, with $\alpha > \frac{1-\epsilon_*}{2}$, for Q-almost all environment ω, and for all large enough k,*

$$\lim_{M \to \infty} E_\omega[\xi_k \mid \xi_M] = 0 \quad P_\omega\text{-almost surely.} \tag{53}$$

Proof of Theorem 1, item 2 By Proposition 2, we can consider a fixed environment ω for which $K = K(\omega) < \infty$. Take $k \geq K$ large. We will consider P_ω^+, and show that (53) implies (52).

We know that ξ_k is a sum of identically distributed variables X_s, $s \in \mathcal{A}(B^k(0))$. By (53), P_ω^+-almost surely, for each such s,

$$\lim_{M \to \infty} P_\omega^+\big(X_s = + \mid \xi_M\big) = \lim_{M \to \infty} \tfrac{1}{2}\big(E_\omega^+[X_s \mid \xi_M] + 1\big)$$
$$= \lim_{M \to \infty} \tfrac{1}{2}\big(E_\omega^+[\xi_k \mid \xi_M] + 1\big) = \tfrac{1}{2}.$$

This implies that for all large enough k, the distribution of ξ_k under $P_\omega^+(\cdot \mid \xi_M)$ converges when $M \to \infty$ to a symmetric distribution. We can show that this extends to any variable X_t as follows. Take k large enough so that $t \notin \mathcal{A}(B^k(0))$, and write

$$E_\omega^+\big[X_t \mid \xi_M\big] = E_\omega^+\big[E_\omega^+[X_t \mid \xi_k] \mid \xi_M\big]. \tag{54}$$

Since $f(x) := E_\omega^+[X_t \mid \xi_k = x]$ is odd, $f(-x) = -f(x)$, and since ξ_k converges in distribution to a symmetric variable, the right-hand side of (54) converges to zero when $M \to \infty$. By dominated convergence, we thus get

$$E_\omega^+[X_t] = \lim_{M \to \infty} E_\omega^+\big[E_\omega^+[X_t \mid \xi_M]\big] = 0.$$

Similarly, $E_\omega^-[X_t] = 0$, and this finishes the proof. $\qquad\square$

4.6 Proofs of Propositions 3 and 4

The sequence $\xi_k \in [-1, 1]$ is Markovian and temporally non-homogeneous; we can nevertheless estimate its transition probabilities with relative precision. Since the

BHS-model considers the pure majority rule φ_{PMR}, its study can be reduced to the sign variables

$$\sigma_k := \begin{cases} +1 & \text{if } \xi_k > 0, \\ 0 & \text{if } \xi_k = 0, \\ -1 & \text{if } \xi_k < 0. \end{cases}$$

Remember that if $B^k(0)$ is good, then

$$n_1(k) < |\mathcal{A}(B^k(0))| < n_2(k), \tag{55}$$

where the leading term in each $n_\#(k)$ is $\beta_k^{1-\epsilon_*}$. But since $h_{k+1} = \beta_k^{-\alpha}$,

$$h_{k+1}^2 n_\#(k) \begin{cases} \nearrow \infty & \text{if } \alpha < \frac{1-\epsilon_*}{2}, \\ \searrow 0 & \text{if } \alpha > \frac{1-\epsilon_*}{2}. \end{cases} \tag{56}$$

We study the sign changes of the sequence ξ_k:

Lemma 5 *Assume $\varphi = \varphi_{PMR}$. Let $\omega \in \{\infty \to k'\}$. For all large $k \geq k'$, the following holds.*

(1) *For all $h_k \searrow 0$,*

$$P_\omega\big(\sigma_k \leq 0 \mid \sigma_{k+1} = +\big) \leq e^{-h_{k+1}^2 n_1(k)/16}.$$

(2) *There exists $c_0 > 0$ such that if $h_k = \beta_{k-1}^{-\alpha}$,*

$$P_\omega\big(\sigma_k \leq 0 \mid \sigma_{k+1} = +\big) \begin{cases} \leq e^{-h_{k+1}^2 n_1(k)/16} & \text{if } \alpha < \frac{1-\epsilon_*}{2}, \\ \geq c_0 & \text{if } \alpha > \frac{1-\epsilon_*}{2}. \end{cases}$$

Proof (1) On $\omega \in \{\infty \to k'\}$, each block $B^k(0)$, $k \geq k'$, is good. In particular, (55) holds and by (35), under $P_\omega\big(\cdot \mid \sigma_{k+1} = +\big)$, the variables $\{X_s, s \in \mathcal{A}(B^k(0))\}$ are i.i.d. with

$$E_\omega[X_s \mid \sigma_{k+1} = +] = h_{k+1}\varphi_{PMR}(\xi_{k+1}) \equiv h_{k+1}.$$

Let $X_s' := X_s - E_\omega[X_s \mid \sigma_{k+1} = +]$. By the Bernstein Inequality,

$$\begin{aligned} & P_\omega(\sigma_k \leq 0 \mid \sigma_{k+1} = +) \\ & = P_\omega\Big(\frac{1}{|\mathcal{A}(B^k(0))|} \sum_{s \in \mathcal{A}(B^k(0))} X_s' \leq -h_{k+1} \ \Big| \ \sigma_{k+1} = +\Big) \\ & \leq e^{-h_{k+1}^2 |\mathcal{A}(B^k(0))|/16} \leq e^{-h_{k+1}^2 n_1(k)/16}. \end{aligned}$$

(2) To prove the lower bound we let $A(k) := |\mathcal{A}(B^k(0))|$ and let $\mathcal{L}_k$ denote the set of integers between 0 and $\sqrt{A(k)}$ that have the same parity as $A(k)$. Using Stirling's formula:

$$
\begin{aligned}
P_\omega\big(\sigma_k \leq 0 \,|\, \sigma_{k+1} = +\big) & \\
&\geq \sum_{L \in \mathcal{L}_k} P_\omega\Big(\sum_{s \in \mathcal{A}(B^k(0))} X_s = -L \,\Big|\, \sigma_{k+1} = + \Big) \\
&= \sum_{L \in \mathcal{L}_k} \binom{A(k)}{\frac{A(k)+L}{2}} \big(\tfrac{1}{2}(1 + h_{k+1})\big)^{\frac{A(k)-L}{2}} \big(\tfrac{1}{2}(1 - h_{k+1})\big)^{\frac{A(k)+L}{2}} \\
&\geq \tilde{c}_0 e^{-h_{k+1}^2 A(k)} e^{-\big(h_{k+1}\sqrt{A(k)}\big)/4}.
\end{aligned}
$$

$\square$

Proof of Proposition 3 Let $\omega \in \{\infty \to k'\}$. The probability we are interested in is defined using the $+$ boundary condition:

$$
P_\omega^+(\xi_{k'} > 0) = \lim_{N \to \infty} P_\omega^{+,N}(\xi_{k'} > 0).
$$

We choose N large, always to be between two successive sets $\mathcal{A}(B^{M-1}(0))$, $\mathcal{A}(B^M(0))$. A lower bound is obtained by assuming that all the sign of the boundary condition travels down to $\xi_{k'}$:
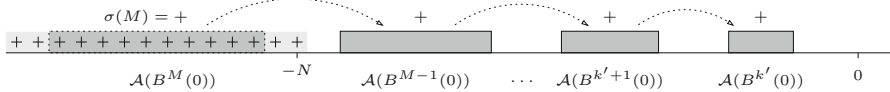
Using Lemma 5,

$$
\begin{aligned}
P_\omega^{+,N}(\xi_{k'} > 0) = P_\omega^{+,N}(\sigma_{k'} = +) & \\
&\geq \prod_{k=k'}^{M-1} P_\omega^{+,N}(\sigma_k = + \,|\, \sigma_{k+1} = +) \\
&\geq \prod_{k=k'}^{\infty} \Big\{ 1 - e^{-h_{k+1}^2 n_1(k)/16} \Big\}.
\end{aligned}
$$

When $\alpha < \frac{1-\epsilon_*}{2}$, we have

$$
\sum_k e^{-h_{k+1}^2 n_1(k)/16} < \infty, \tag{57}
$$

and so that last product converges and goes to 1 when $k' \to \infty$. $\square$

Proof of Proposition 4 By Proposition 2, we can consider a fixed environment ω for which $K = K(\omega) < \infty$. Then for each $k' > K$, we have that $\omega \in \{\infty \to k'\}$.

Take $M > k'$. If $\xi_M = 0$ then $E_\omega[\xi_{k'} \mid \xi_M] = 0$. If $\xi_M > 0$, then by the attractiveness of the model, $E_\omega[\xi_{k'} \mid \xi_M] \geq 0$. For an upper bound, we look for the scale k at which ξ changes sign:

$$\kappa_M := \max\{k' \leq k \leq M : \xi_k \leq 0\}, \tag{58}$$

with the convention that $\kappa_M := k' - 1$ if $\xi_k > 0$ for all $k' \leq k < M$. Again by attractiveness, on $\{\kappa_M \geq k'\}$, the Markov property gives $E_\omega[\xi_{k'} \mid \xi_{\kappa_M}] \leq 0$. Therefore, $E_\omega[\xi_{k'} \mid \kappa_M \geq k', \xi_M] \leq 0$, and so

$$E_\omega[\xi_{k'} \mid \xi_M] \leq P_\omega(\kappa_M = k' - 1 \mid \xi_M). \tag{59}$$

It follows by Lemma 5 and $1 - x \leq e^{-x}$ that when $\alpha > \frac{1-\epsilon_*}{2}$,

$$\begin{aligned}
P_\omega(\kappa_M = k' - 1 \mid \xi_M) &= P_\omega(\xi_{k'} > 0, \ldots, \xi_{M-1} > 0 \mid \xi_M) \\
&= \prod_{k=k'}^{M-1} P_\omega(\sigma_k = + \mid \sigma_{k+1} = +) \\
&\leq (1 - c_0)^{M-k'}.
\end{aligned}$$

$\square$

4.7 Proof of Theorem 2

The proof of uniqueness when φ is Lipshitz at the origin will be based on the same principle used when proving item 2 of Theorem 1.

Proposition 5 *Let $\mathbb{P} = Q \otimes P_\omega$ be the distribution of any process specified by g. Under the hypotheses of Theorem 2, for Q-almost all environment ω, and for all large enough k',*

$$\lim_{M \to \infty} E_\omega[\xi_{k'} \mid \xi_M] = 0 \quad P_\omega\text{-almost surely.}$$

We consider an environment ω with $K = K(\omega) < \infty$. We take $k' > K$, and $M > k'$. As before, the proof is based on showing that whatever the sign of ξ_M, the sequence ξ_k has a positive probability of having changed sign before reaching k'.

We will look at the variables ξ_k at even times: $\xi_M, \xi_{M-2}, \ldots$, and show that the probability of ξ changing sign between two scales k and $k - 2$ is bounded away from zero.

Lemma 6 *Let $\omega \in \{\infty \to k'\}$ with k' large enough. If φ satisfies (15), then for all $\alpha > 0$, and all $k \geq k' + 2$,*

$$P_\omega(\xi_{k-2} \leq 0 \mid \xi_k) \geq \frac{c_1}{2} > 0,$$

where c_1 is a universal constant.

Proof Let $\gamma > 0$ be such that (15) holds. If $\xi_k \leq 0$, then attractiveness gives $P_\omega(\xi_{k-2} \leq 0 \,|\, \xi_k) \geq \frac{1}{2}$. If $\xi_k = x > 0$, let us denote $P_\omega^x(\cdot) := P_\omega(\cdot \,|\, \xi_k = x)$. We have

$$
\begin{aligned}
P_\omega^x(\xi_{k-2} \leq 0) &= E_\omega^x\big[P_\omega^x(\xi_{k-2} \leq 0 \,|\, \xi_{k-1})\big] \\
&\geq E_\omega^x\big[P_\omega^x(\xi_{k-2} \leq 0 \,|\, \xi_{k-1}) 1_{\{|\xi_{k-1}| \leq h_{k-1}^m\}}\big],
\end{aligned}
$$

where $m > 0$ is chosen such that

$$
\alpha > \frac{1 - \epsilon_*}{2(1 + m\gamma)}. \tag{60}
$$

Again, by attractiveness, $-h_{k-1}^m \leq \xi_{k-1} \leq 0$ implies

$$
P_\omega^x(\xi_{k-2} \leq 0 \,|\, \xi_{k-1}) \geq \tfrac{1}{2}.
$$

When $0 \leq \xi_{k-1} \leq h_{k-1}^m$, we let

$$
\overline{X}_s := \frac{X_s - E_\omega^x[X_s | \xi_{k-1}]}{\sqrt{\mathrm{Var}_\omega^x(X_s \,|\, \xi_{k-1})}},
$$

which are centered with variance 1. Then

$$
\begin{aligned}
&P_\omega^x(\xi_{k-2} \leq 0 \,|\, \xi_{k-1}) \\
&= P_\omega^x\left(\frac{1}{\sqrt{|\mathcal{A}(B^{k-2}(0))|}} \sum_{s \in \mathcal{A}(B^{k-2}(0))} \overline{X}_s \leq -\frac{E_\omega^x[X_s \,|\, \xi_{k-1}]\sqrt{|\mathcal{A}(B^{k-2}(0))|}}{\sqrt{\mathrm{Var}_\omega^x(X_s \,|\, \xi_{k-1})}} \,\bigg|\, \xi_{k-1} \right).
\end{aligned}
$$

By (15), there exists $0 < \lambda < \infty$ and $\delta > 0$ such that $\varphi(y) \leq \lambda y^\gamma$ for all $0 \leq y \leq \delta$. Therefore, if k is large enough so that $h_{k-1}^m \leq \delta$,

$$
0 \leq E_\omega^x[X_s \,|\, \xi_{k-1}] = h_{k-1}\varphi(\xi_{k-1}) \leq \lambda h_{k-1}\xi_{k-1}^\gamma \leq \lambda h_{k-1}^{1+m\gamma}.
$$

Then, since $B^{k-2}(0)$ is good,

$$
\frac{E_\omega^x[X_s \,|\, \xi_{k-1}]\sqrt{|\mathcal{A}(B^{k-2}(0))|}}{\mathrm{Var}_\omega^x(X_s \,|\, \xi_{k-1})} \leq \frac{\lambda h_{k-1}^{1+m\gamma}\sqrt{n_2(k-2)}}{\sqrt{1 - h_{k-1}^2}}.
$$

But, the dominating term in this last expression is $\beta_{k-2}^{-(\alpha(1+m\gamma)-(1-\epsilon_*)/2)}$, which tends to zero since (60) holds. Therefore, taking k large enough,

$$
\begin{aligned}
P_\omega^x(\xi_{k-2} \leq 0 \,|\, \xi_{k-1}) &\geq P_\omega^x\left(\frac{1}{\sqrt{|\mathcal{A}(B^{k-2}(0))|}} \sum_{s \in \mathcal{A}(B^{k-2}(0))} \overline{X}_s \leq -1 \,\bigg|\, \xi_{k-1} \right) \\
&\geq 0,5 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1} e^{-x^2/2}\, dx \equiv c_1.
\end{aligned}
$$

It remains to study $P_\omega^x(|\xi_{k-1}| \le h_{k-1}^m)$. We have $E_\omega^x[\xi_{k-1}] = h_k\varphi(x)$, so if k is such that $h_k \le h_{k-1}^m/2$ then, using Chebyshev's inequality,

$$P_\omega^x(|\xi_{k-1}| \le h_{k-1}^m) \ge P_\omega^x(|\xi_{k-1} - h_k\varphi(x)| \le h_{k-1}^m/2)$$
$$\ge 1 - \frac{2}{|\mathcal{A}(B^{k-1}(0))|h_{k-1}^m} \ge 1 - \frac{2}{n_1(k-1)h_{k-1}^m},$$

which is $\ge \frac{1}{2}$ when k is large enough. $\qquad\square$

Proof of Proposition 5 The proof is the same as the one of Proposition 4. If $\xi_M > 0$, define S_M as in (58). Assuming for simplicity that $M - k'$ is even, Lemma 6 gives

$$P_\omega(S_M = k' - 1 \mid \xi_M) \le P_\omega(\xi_{k'} > 0, \xi_{k'+2} > 0, \dots, \xi_{M-2} > 0 \mid \xi_M)$$
$$\le \left(1 - \frac{c_1}{2}\right)^{(M-k')/2}. \qquad\square$$

Remark 6 Proceeding as in the proof of Proposition 3, we show that it is possible to obtain non-uniqueness with a φ continuous at the origin, and the sequence h_k as in (14) with $\alpha < \frac{1-\epsilon_*}{2}$ (as mentioned after Theorem 1, $\alpha > \frac{1-\epsilon_*}{2}$ always leads to uniqueness, for any function φ). Namely, fix $\omega \in \{\infty \to k'\}$. We want to define a sequence $\delta_k \searrow 0$ so that the probability

$$P_\omega^{+,N}(\xi_{k'} > 0) \ge P_\omega^{+,N}(\xi_{k'} \ge \delta_{k'}, \dots, \xi_M \ge \delta_M)$$
$$\ge \prod_{k=k'}^{M-1} P_\omega^{+,N}(\xi_k \ge \delta_k \mid \xi_{k+1} = \delta_{k+1})$$

is close to one. The majority rule φ will be defined appropriately along the sequence δ_k, then extended to $[0, 1]$ by linear interpolation, so as to be continuous at the origin and to satisfy $\varphi(0) = 0$. Finally, φ is defined in $[-1, 0]$ in order to be odd. Conditioned on $\{\xi_{k+1} = \delta_{k+1}\}$, the variables $X_s, s \in \mathcal{A}(B^k(0))$ are Bernoulli with expectation

$$E_\omega[X_s \mid \xi_{k+1} = \delta_{k+1}] = h_{k+1}\varphi(\delta_{k+1}) \equiv \tau_k.$$

If $\delta_k < \tau_k$, for all k, then $\in \mathcal{A}(B^k(0))$, then

$$P_\omega^{+,N}(\xi_k \ge \delta_k \mid \xi_{k+1} = \delta_{k+1}) \ge P_\omega^{+,N}(\delta_k \le \xi_k \le (2\tau_k - \delta_k) \mid \xi_{k+1} = \delta_{k+1})$$
$$= P_\omega^{+,N}(|\xi_k - \tau_k| \le (\tau_k - \delta_k) \mid \xi_{k+1} = \delta_{k+1}).$$

Using Chebyshev's inequality,

$$P_\omega^{+,N}(|\xi_k - \tau_k| \le (\tau_k - \delta_k) \mid \xi_{k+1} = \delta_{k+1}) \ge 1 - \frac{1}{n_1(k)(\tau_k - \delta_k)^2}.$$

By taking for example

$$\delta_k := \beta_{k+1}^{-\alpha}, \quad \varphi(\beta_{k+1}^{-\alpha}) := \beta_{k-2}^{-\alpha},$$

we get, since $\alpha < \frac{1-\epsilon_*}{2}$,

$$P_\omega^{+,N}\left(\xi_{k'} > 0\right) \geq \prod_{k=k'}^{M-1} \left\{ 1 - 2\beta_k^{-((1-\epsilon_*)-2\alpha)} \beta_{k-1}^{2\alpha} \right\}$$

$$\geq \exp\left\{ -\sum_{k=k'}^{\infty} \beta_k^{-((1-\epsilon_*)-2\alpha)} \beta_{k-1}^{2\alpha} \right\}.$$

As a consequence,

$$\lim_{k' \to \infty} P_\omega^+\left(\xi_{k'} > 0\right) = 1.$$

Of course, in contrast with (15), the majority rule constructed above satisfies

$$\limsup_{z \to 0^+} \frac{\varphi(z)}{z^\gamma} = \infty, \qquad \text{for all } \gamma > 0.$$

5 Proof of Theorem 3

The proof of uniqueness when φ is Lipschitz, for arbitrary h_ks, will be based on the same principle used when proving item 2 of Theorem 1, showing that for all large enough k, the distribution of ξ_k under $P_\omega(\cdot \mid \xi_M)$ converges to a symmetric distribution when $M \to \infty$:

Proposition 6 *Assume φ is Lipshitz in a neighborhood of the origin. Let $h_k \searrow 0$ be an arbitrary sequence. Let $\mathbb{P} = Q \otimes P_\omega$ be the distribution of any process specified by g. Then for Q-almost all environment ω, for all large enough k',*

$$\lim_{M \to \infty} E_\omega\left[\xi_{k'} \mid \xi_M\right] = 0 \quad P_\omega\text{-almost surely.}$$

To understand why Lipschitzness near the origin implies uniqueness regardless of the details of the sequence h_k, we first consider a particular case.

Assume φ is globally linear with slope 1:

$$\varphi_{ID}(z) := z \quad \forall z \in [-1, 1].$$

Let $\omega \in \{\infty \to k'\}$, and take $k > k'$. Then

$$E_\omega[\xi_k] = E_\omega\left[E_\omega[\xi_k \mid \xi_{k+1}]\right] = E_\omega\left[h_{k+1}\varphi_{ID}(\xi_{k+1})\right] = h_{k+1} E_\omega[\xi_{k+1}].$$

Repeating this procedure we get, for all $L \geq 1$,

$$E_\omega[\xi_k] = \left\{ \prod_{j=k+1}^{k+L} h_j \right\} E_\omega[\xi_{k+L}]. \tag{61}$$

Taking $L \to \infty$ gives $E_\omega[\xi_k] = 0$.

The proof of Proposition 6 consists in using this phenomenon, which obviously doesn't depend on the precise values of the sequence h_k. So first, we will consider a case where the Lipschitzness of φ is global:

Lemma 7 *Assume that $\widetilde{\varphi}$ is 1-Lipschitz on $[-1, 1]$:*

$$|\widetilde{\varphi}(z_2) - \widetilde{\varphi}(z_1)| \leq |z_2 - z_1| \quad \forall z_1, z_2 \in [-1, 1].$$

Let $\mathbb{P} = Q \otimes P_\omega$ be the distribution of any process specified by the g-function associated to $\widetilde{\varphi}$ and to some sequence $\tilde{h}_k \searrow 0$. Then for Q-almost all environment ω, for all large enough k',

$$\lim_{M \to \infty} E_\omega[\tilde{\xi}_{k'} \mid \tilde{\xi}_M] = 0 \quad P_\omega\text{-almost surely.}$$

Proof of Proposition 6: Let $\delta > 0$ and $\lambda > 0$ be such that $0 \leq \varphi(z_2) - \varphi(z_1) \leq \lambda(z_2 - z_1)$ for all $-\delta \leq z_1 \leq z_2 \leq \delta$. We define a function $\widetilde{\varphi}$ that satisfies the conditions of Lemma 7:

$$\widetilde{\varphi}(z) := \tfrac{1}{\lambda}\phi(z),$$

where

$$\phi(z) := \begin{cases} \lambda(z + \delta) + \varphi(-\delta) & \text{if } z \in [-1, -\delta], \\ \varphi(z) & \text{if } z \in [-\delta, \delta], \\ \lambda(z - \delta) + \varphi(\delta) & \text{if } z \in [\delta, 1]. \end{cases} \tag{62}$$

Assume $\omega \in \{\infty \to k'\}$ for some large k'. We fix $M > k'$ large, and to study $E_\omega[\xi_{k'} \mid \xi_M]$ we construct the sequence ξ_k for k decreasing from M to k', coupled to another sequence $\tilde{\xi}_k$; ξ_k will have its transition probability fixed by φ and h_k, and $\tilde{\xi}_k$ will have its transition probability fixed by $\widetilde{\varphi}$ and

$$\tilde{h}_k := \lambda h_k.$$

If λ is large, we may need to take k' large enough so that $\tilde{h}_k \leq 1/2$ for all $k \geq k'$. The processes ξ_k and $\tilde{\xi}_k$ will be constructed along with a sequence $\gamma_k \in \{0, 1\}$: if $\gamma_k = 1$, then ξ_k and $\tilde{\xi}_k$ are still coupled; $\gamma_k = 0$ means they have already decoupled.

For simplicity, we will continue denoting the coupling measure by P_ω. The construction is illustrated on the figure below. We start with some fixed ξ_M. By (51), we can assume that M is large enough in order to guarantee that $|\xi_M| \leq \delta$. Let then $\tilde{\xi}_M := \xi_M$, and $\gamma_M := 1$.

Given $(\gamma_{k+1}, \tilde{\xi}_{k+1}, \xi_{k+1})$, $(\gamma_k, \tilde{\xi}_k, \xi_k)$ is constructed as follows:

(1) Sample $\tilde{\xi}_k$ as an average of variables $\{\widetilde{X}_s = \pm, s \in \mathcal{A}(B^k(0))\}$, i.i.d. with

$$E_\omega[\widetilde{X}_s \,|\, \widetilde{\xi}_{k+1}] = \widetilde{h}_{k+1}\widetilde{\varphi}(\widetilde{\xi}_{k+1}).$$

(2) If $\gamma_{k+1} = 1$, set $\xi_k := \widetilde{\xi}_k$ and $\gamma_k := 1_{\{|\xi_k| \le \delta\}}$.

(3) If $\gamma_{k+1} = 0$, set $\gamma_k := 0$, and sample ξ_k as an average of variables $\{X_s = \pm, s \in \mathcal{A}(B^k(0))\}$, i.i.d. with

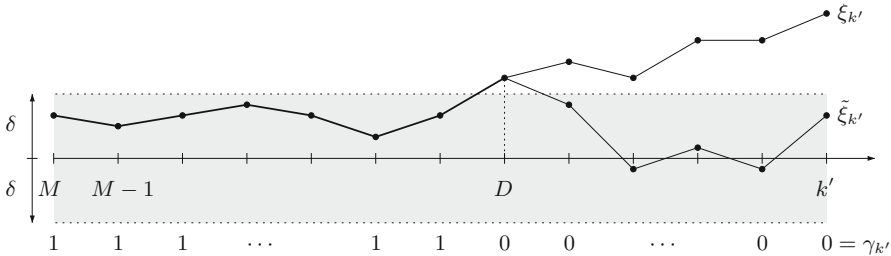$$E_\omega[X_s \,|\, \xi_{k+1}] = h_{k+1}\varphi(\xi_{k+1}).$$

By construction, $(\gamma_k, \widetilde{\xi}_k, \xi_k)$ is a Markov chain, and ξ_k has the proper marginals. Namely, if $\gamma_{k+1} = 0$ then

$$E_\omega[\xi_k \,|\, \xi_{k+1}] = h_{k+1}\varphi(\xi_{k+1}),$$

and if $\gamma_{k+1} = 1$, then

$$
\begin{aligned}
E_\omega[\xi_k \,|\, \xi_{k+1}] = E_\omega[\widetilde{\xi}_k \,|\, \widetilde{\xi}_{k+1}] &= \widetilde{h}_{k+1}\widetilde{\varphi}(\widetilde{\xi}_{k+1}) \\
&= \widetilde{h}_{k+1}\widetilde{\varphi}(\xi_{k+1}) \\
&= h_{k+1}\varphi(\xi_{k+1}),
\end{aligned}
$$

where in the last line we used that $|\xi_{k+1}| \le \delta$.

Let D be the scale at which decoupling occurs:

$$D := \max\{k : M > k \ge k', \gamma_k = 0\}.$$

Clearly, D is a stopping time for the chain $(\gamma_k, \widetilde{\xi}_k, \xi_k)$, and if $k \in \{k : M \ge k > D\}$, then $\xi_k = \widetilde{\xi}_k$ and $|\xi_k| = |\widetilde{\xi}_k| \le \delta$.

Fix $k' < L < M$. On the one hand, proceeding as in (50),

$$
\begin{aligned}
\left|E_\omega[\xi_{k'}, D \ge L \,|\, \xi_M]\right| &\le P_\omega(D \ge L \,|\, \xi_M) \\
&\le \sum_{k \ge L} P_\omega(|\xi_k| \ge \delta \,|\, \xi_M) \le \sum_{k \ge L} e^{-cn_1(k)}.
\end{aligned}
$$

On the other hand, $D < L$ implies $\xi_L = \widetilde{\xi}_L$ and so

$$E_\omega[\xi_{k'}, D < L \,|\, \xi_M] = E_\omega\left[E_\omega[\xi_{k'} \,|\, \xi_L]1_{\{D<L\}} \,\Big|\, \xi_M\right]$$

$$= E_\omega\Big[f_\omega(\tilde\xi_L) 1_{\{D<L\}} \,\Big|\, \xi_M \Big]$$
$$= E_\omega\big[f_\omega(\tilde\xi_L) \,\big|\, \xi_M \big] + O\big(P_\omega(D \geq L \,|\, \xi_M)\big),$$

where $f_\omega(x) := E_\omega[\xi_{k'} \,|\, \xi_L = x]$. Now, the construction of $\tilde\xi$ was based on $\widetilde\varphi$, so by Lemma 7,

$$E_\omega[\tilde\xi_L \,|\, \tilde\xi_M] \to 0 \quad \text{when } M \to \infty.$$

Since $\tilde\xi_M := \xi_M$, $\tilde\xi_L$ has, under $P_\omega(\cdot \,|\, \xi_M)$ in the limit $M \to \infty$, a symmetric law. But since $f_\omega(-x) = -f_\omega(x)$, this implies
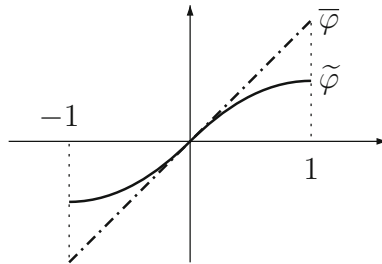
$$E_\omega\big[f_\omega(\tilde\xi_L) \,\big|\, \xi_M \big] \to 0 \quad \text{as } M \to \infty.$$

We have thus shown that for all $L > k'$,

$$\limsup_{M\to\infty} \big| E_\omega[\xi_{k'} \,|\, \xi_M] \big| \leq 2 \sum_{k \geq L} e^{-cn_1(k)}.$$

$$\square$$

Proof of Lemma 7 We work with two different g-functions that have the same sequence $\tilde h_k$ but different majority rules. The first, $\widetilde g$, is associated to $\widetilde\varphi$, which is 1-Lipschitz $[-1, 1]$. The second, $\overline g$, is associated to $\overline\varphi := \varphi_{ID}$:

Uniformly in $z_1 < z_2$,

$$0 \leq \widetilde\varphi(z_2) - \widetilde\varphi(z_1) \leq z_2 - z_1 \equiv \overline\varphi(z_2) - \overline\varphi(z_1). \tag{63}$$

Let $\widetilde X$ (resp. $\overline X$) denote the process associated to $\widetilde g$ (resp. $\overline g$). Using attractiveness and the notations of Sect. 4.2,

$$E_\omega^{-,N}[\tilde\xi_{k'}] \leq E_\omega[\tilde\xi_{k'} \,|\, \tilde\xi_M] \leq E_\omega^{+,N}[\tilde\xi_{k'}],$$

where N is chosen appropriately in function of M. Since $\tilde\xi_{k'}$ is an average of identically distributed variables $\widetilde X_s$, our aim will be to show that when $M \to \infty$,

$$0 \leq E_\omega^{+,N}[\tilde\xi_{k'}] - E_\omega^{-,N}[\tilde\xi_{k'}] = E_\omega^{+,N}[\widetilde X_s] - E_\omega^{-,N}[\widetilde X_s] \to 0.$$

To bound this last difference, consider the coupling of $E_\omega^{+,N}$ and $E_\omega^{-,N}$ described in Sect. 4.2, which we here denote by $\mathsf{E}_\omega^{\pm,N}$. Since that coupling is maximal,

$$E_\omega^{+,N}[\widetilde{X}_s] - E_\omega^{-,N}[\widetilde{X}_s] = 2\mathsf{E}_\omega^{\pm,N}[1_{\{\widetilde{\Delta}_s=\binom{+}{-}\}}]. \tag{64}$$

We will now use (63) to further couple the pair processes, $\widetilde{\Delta}_s = \binom{\widetilde{X}_s^2}{\widetilde{X}_s^1}$ and $\overline{\Delta}_s = \binom{\overline{X}_s^2}{\overline{X}_s^1}$. This coupling will contain the four processes associated to $\widetilde{g}$ and $\overline{g}$, with boundary conditions $+$ and $-$. The coupling will be such that there are more discrepancies between $\overline{X}^2$ and $\overline{X}^1$ than there are between $\widetilde{X}^2$ and $\widetilde{X}^1$, in the following sense:

$$1_{\{\widetilde{\Delta}_s=\binom{+}{-}\}} \leq 1_{\{\overline{\Delta}_s=\binom{+}{-}\}} \quad a.s. \tag{65}$$

By definition, when $s \leq -N$, $\widetilde{\Delta}_s = \overline{\Delta}_s = \binom{+}{-}$. Let U_t, $t > -N$ be an i.i.d. sequence, each with U_t uniform on $[0, 1]$. Assume all pairs, $\widetilde{\Delta}_s = \binom{\widetilde{x}^2}{\widetilde{x}^1}$ and $\overline{\Delta}_s = \binom{\overline{x}^2}{\overline{x}^1}$, have been sampled for all $s < t$ and that (65) holds for all $s < t$. Let $\widetilde{\Delta}_t$ (resp. $\overline{\Delta}_t$) be defined as in (41), in which A_t, B_t, C_t are replaced by the corresponding $\widetilde{A}_t, \widetilde{B}_t, \widetilde{C}_t$ (resp. $\overline{A}_t, \overline{B}_t, \overline{C}_t$). (Obs: we are using the same variable U_t to define $\widetilde{\Delta}_t$ and $\overline{\Delta}_s$.) Then $\widetilde{\Delta}_t$ and $\overline{\Delta}_t$ obviously have the correct distribution. To verify that (65) holds at time t, we first remind that

$$\overline{A}_t = \{0 \leq U_t < \overline{g}_t^\omega(+ \,|\, (\overline{x}^2)_{-\infty}^{t-1}) - \overline{g}_t^\omega(+ \,|\, (\overline{x}^1)_{-\infty}^{t-1})\},$$
$$\widetilde{A}_t = \{0 \leq U_t < \widetilde{g}_t^\omega(+ \,|\, (\widetilde{x}^2)_{-\infty}^{t-1}) - \widetilde{g}_t^\omega(+ \,|\, (\widetilde{x}^1)_{-\infty}^{t-1})\}.$$

Using the fact that $\overline{X}$ has more discrepancies than $\widetilde{X}$, and (63),

$$\overline{g}_t^\omega(+ \,|\, (\overline{x}^2)_{-\infty}^{t-1}) - \overline{g}_t^\omega(+ \,|\, (\overline{x}^1)_{-\infty}^{t-1}) = \frac{\widetilde{h}_{k_t}}{2}\left\{\overline{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\overline{x}_s^2\right) - \overline{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\overline{x}_s^1\right)\right\}$$

$$\geq \frac{\widetilde{h}_{k_t}}{2}\left\{\overline{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\widetilde{x}_s^2\right) - \overline{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\widetilde{x}_s^1\right)\right\}$$

$$\geq \frac{\widetilde{h}_{k_t}}{2}\left\{\widetilde{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\widetilde{x}_s^2\right) - \widetilde{\varphi}\left(\frac{1}{|S_t|}\sum_{s\in S_t}\widetilde{x}_s^1\right)\right\}$$

$$= \widetilde{g}_t^\omega(+ \,|\, (\widetilde{x}^2)_{-\infty}^{t-1}) - \widetilde{g}_t^\omega(+ \,|\, (\widetilde{x}^1)_{-\infty}^{t-1}),$$

which implies $\widetilde{A}_t \subset \overline{A}_t$ almost surely.

With (65) at hand, we go back to (64):

$$E_\omega^{+,N}[\widetilde{X}_s] - E_\omega^{-,N}[\widetilde{X}_s] = 2\mathsf{E}_\omega^{\pm,N}[1_{\{\widetilde{\Delta}_s=\binom{+}{-}\}}]$$
$$\leq 2\mathsf{E}_\omega^{\pm,N}[1_{\{\overline{\Delta}_s=\binom{+}{-}\}}] = E_\omega^{+,N}[\overline{X}_s] - E_\omega^{-,N}[\overline{X}_s].$$

But since $\overline{\varphi}$ is purely linear, the explicit computation made at the beginning of the section can be repeated, giving

$$E_\omega^{\pm, N}[\overline{X}_s] = \left\{ \prod_{k=k'+1}^{M} \tilde{h}_k \right\} (\pm 1) \to 0 \quad \text{when } M \to \infty.$$

□

6 Concluding remarks

The analysis of the model was possible due to the Markovian structure of the sequence ξ_k, in particular to the relation (valid on $\{\infty \to k\}$)

$$E_\omega[\xi_k] = h_{k+1} E_\omega[\varphi(\xi_{k+1})]. \tag{66}$$

We will give a simple heuristic argument that might shed some light on the proofs given above, and on the role played by the continuity of φ at the origin.

A mean field approximation consists in assuming that ξ_{k+1} can be approximated by its mean:

$$\xi_{k+1} \simeq E_\omega[\xi_{k+1}].$$

This allows to transform $E_\omega[\varphi(\xi_{k+1})] \simeq \varphi(E_\omega[\xi_{k+1}])$. This approximation is correct in exactly one case: when φ is purely linear.

With the mean field approximation, one can transform (66) into a deterministic toy model, in which $\mu_k := E_\omega[\xi_k]$ is a sequence satisfying the relation

$$\mu_k = h_{k+1} \varphi(\mu_{k+1}). \tag{67}$$

We thus take some large integer M, fix some initial condition, μ_M, and study the sequence $\mu_M, \mu_{M-1}, \ldots, \mu_{k_*}$. Since $\varphi(0) = 0$, 0 is always a fixed point for the dynamics. In the case of a purely linear majority rule, $\varphi(z) = \lambda z$, the trajectory of $\mu_{k+1} \to \mu_k$ is always attracted towards the origin, independently of the initial condition. For example, if $\mu_M > 0$:
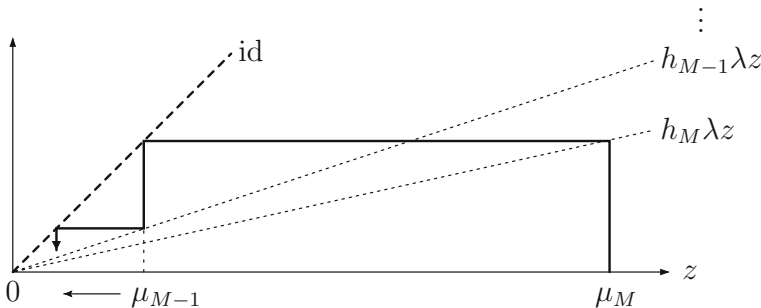
The same qualitative behavior holds if $\varphi'(0) < \infty$. Namely, when φ is Lipschitz near the origin, a coupling with a straight line has shown that the same phenomenon occurs: for large enough k, $z \geq 0$, the curve $z \mapsto h_{k+1}\varphi(z)$ lies strictly *below* the identity $z \mapsto z$, and any initial condition is also attracted towards the origin.

In the case of a pure majority rule [or when $\varphi'(0) = \infty$ as in Remark 6], the mechanism changes: the trajectories of $\mu_{k+1} \to \mu_k = h_{k+1}\varphi_{PMR}(\mu_{k+1})$, are repelled away from the origin, with a sign that depends on the sign of the initial condition. For example, if $\mu_M > 0$:
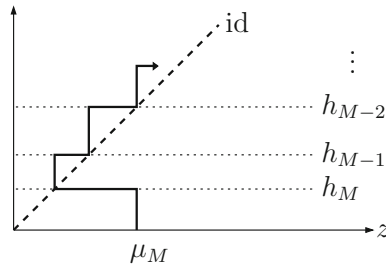
This scenario was shown to hold for the BHS model, at least when α was taken small enough. When α is large, fluctuations allow μ_k to change sign at any time, yielding uniqueness. The role of $\varphi'(0)$ in the model of Bramson and Kalikow is under current investigation (work in progress).

Appendix A. The uniqueness criterion

The uniqueness criterion (52) is standard in attractive systems, although usually used for translation invariant processes (which is not the case of P_ω). See for example how it is used in Statistical Mechanics in [10], or in [7] for g-measures.

We will use some notations and results from Sect. 4.2. Let $\chi := \{\pm\}^{\mathbb{Z}} = \{x = (x_t)_{t \in \mathbb{Z}}, x_t = \pm\}$ equipped with the σ-field generated by cylinders. A function $f : \chi \to \mathbb{R}$ is local if it depends only on a finite number of x_ts; we denote its support by $\mathrm{supp}(f)$. We say f is increasing if $f(x) \leq f(y)$ whenever $x_t \leq y_t$ for all t.

The probability measures on χ are entirely determined by their action on local functions: if $E_1[f] = E_2[f]$ for all local function f, then $P_1 = P_2$. But a local function can always be represented as

$$f(x) = \sum_{B \subset \mathrm{supp}(f)} \alpha_B n_B,$$

where $\alpha_B \in \mathbb{R}$, $n_B := \prod_{t \in B} n_t$, with $n_t := \frac{1}{2}(1 + x_t)$.

Lemma 8 *If $E_\omega^+[X_t] = E_\omega^-[X_t]$ for all $t \in \mathbb{Z}$, then $P_\omega^+ = P_\omega^-$.*

Proof By what was said above, it suffices to show that $E_\omega^+[n_B] = E_\omega^-[n_B]$ for all finite $B \subset \mathbb{Z}$. Observe that $G = \sum_{t \in B} n_t - n_B$ is increasing. Using (43) yields $E_\omega^+[G] \geq E_\omega^-[G]$, which can be written

$$\sum_{t \in B} \left(E_\omega^+[n_t] - E_\omega^-[n_t] \right) \geq E_\omega^+[n_B] - E_\omega^-[n_B] \geq 0.$$

But $E_\omega^+[n_t] - E_\omega^-[n_t] = E_\omega^+[X_t] - E_\omega^-[X_t] = 0$. □

Lemma 9 *If $P_\omega^+ = P_\omega^-$, then any other measure P_ω satisfying* (3) *coincides with P_ω^+ and P_ω^-.*

Proof Once again, we need only consider local functions of the form n_B. If $B \subset [a, b]$, we can take k large enough so that $\mathcal{A}(B^k(0)) \subset (-\infty, a \wedge 0)$. Then,

$$E_\omega[n_B] = E_\omega \big[E_\omega[n_B \mid \xi_k] \big].$$

But, since n_B is increasing,

$$E_\omega[n_B \mid \xi_k] \leq E_\omega[n_B \mid \xi_k = +1] \equiv E_\omega^{+, N(k)}[n_B],$$

for some suitable $N(k)$. But by the definition of P_ω^+,

$$\lim_{k \to \infty} E_\omega^{+, N(k)}[n_B] = E_\omega^+[n_B].$$

We therefore have $E_\omega[n_B] \leq E_\omega^+[n_B]$. In the same, way, $E_\omega^-[n_B] \leq E_\omega[n_B]$. As a consequence, $E_\omega[n_B] = E_\omega^+[n_B] = E_\omega^-[n_B]$, and the same extends to all local function f. □

References

1. Berger, N., Hoffman, C., Sidoravicius, V.: Non-uniqueness for specifications in $\ell^{2+\epsilon}$ (preprint) (2005). arXiv:0312344[mat.PR]
2. Bramson, M., Kalikow, S.: Nonuniqueness in g-functions. Isr. J. Math. **84**(1–2), 153–160 (1993)
3. Doeblin, W., Fortet, R.: Sur des chaînes à liaisons complètes. Bull. Soc. Math. France **65**, 132–148 (1937)
4. Fernández, R., Maillard, G.: Uniqueness and non-uniqueness of chains on half lines (preprint) (2007). arXiv:0705.0808[mat.PR]
5. Friedli, S.: A note on the Bramson-Kalikow process. Braz. J. Probab. Stat. (2015) (to appear)
6. Gallesco, C.F., Gallo, S., Takahashi, D.Y.: Explicit estimates in the Bramson-Kalikow model. Nonlinearity **27**, 2281–2296 (2014)
7. Hulse, P.: Uniqueness and ergodic properties of attractive g-measures. Ergod. Theory Dyn. Syst. **11**, 65–77 (1991)
8. Hulse, P.: An example of non-unique g-measures. Ergod. Theory Dyn. Syst. **26**, 439–445 (2006)
9. Johansson, A., Öberg, A.: Square summability of variations of g-functions and uniqueness of g-measures. Math. Res. Lett. **10**(5–6), 587–601 (2003)
10. Preston, C.: Random Fields. Springer, New York (1976)
11. Williams, D.: Probability with Martingales. Cambridge Mathematical Textbooks, Cambridge (1991)