Bit flipping and time to recover

Anton Muratov* Sergei Zuyev*

November 18, 2018

Abstract

We call 'bits' a sequence of devices indexed by positive integers, where every device can be in two states: 0 (idle) and 1 (active). Start from the 'ground state' of the system when all bits are in 0-state. In our first Binary Flipping (BF) model, the evolution of the system is the following: at each time step choose one bit from a given distribution $\mathcal{P}$ on the integers independently of anything else, then flip the state of this bit to the opposite. In our second Damaged Bits (DB) model a 'damaged' state is added: each selected idling bit changes to active, but selecting an active bit changes its state to damaged in which it then stays forever.

In both models we analyse the recurrence of the system's ground state when no bits are active. We present sufficient conditions for both BF and DB models to show recurrent or transient behaviour, depending on the properties of $\mathcal{P}$. We provide a bound for fractional moments of the return time to the ground state for the BF model, and prove a Central Limit Theorem for the number of active bits for both models.

Keywords: binary system; bit flipping; random walk on a countable group; Markov chain recurrence; critical behaviour

AMS 2010 Subject Classification: Primary 60J27; Secondary 60J10; 68Q87

*Chalmers University of Technology, Department of Mathematical Sciences, SE-412 96 Gothenburg, Sweden. Email: `[muratov|sergei.zuyev]@chalmers.se`

1 Introduction and Model Description

In many areas of engineering and science one faces an array of devices which possess a few states. In the simplest case these could be on-off or idle-active states, in other situations a damaged state is also possible. By the analogy with computer science, such a two-state device can be called a *bit* which in some case can also be 'damaged'. If the activation-deactivation cycles (flipping) or damage produce themselves in a random fashion, a natural question to ask is when, if at all, the system of bits recovers to some initial or *ground state* when none of the bits are active, allowing only for idling and damaged bits to be seen. The time to recover may be finite, but, in general, may also assume infinite values when the system actually does not recover. In the latter case we speak of *transient* behaviour of the system. In the former case, depending on whether the mean of the recover time exists or not, we speak of a *positive-* or of a *null-recurrence.* Similarly to random walk models, this classification is tightly related to the exact random mechanism governing the change of the bits' states.

In the present paper we consider two basic models. In both models we deal with a countably infinite array of bits which we index by the positive integers $\mathbb{N} = \{1, 2, \dots\}$. Initially, at step 0, the system is in the ground state, i.e. all the bits are idling. At each next step the index of the bit to change its state is sampled independently of the current state of the bits from a given probability distribution on $\mathbb{N}$,

$$\mathcal{P} = (p_1, p_2, \dots) : \sum_{i=1}^{\infty} p_i = 1.$$

Without loss of generality, we may assume that the bits are indexed in such a way that

$$p_1 \geq p_2 \geq p_3 \geq \cdots$$

so that the bits most likely to change their state are put first. We also assume that the support of the distribution $\mathcal{P}$ is unbounded, otherwise our models are described by a finite state Markov chain with an evident behaviour. The main quantities of interest are the number of steps τ until the first return to the ground state, and η_n – the number of bits being active at step n. The two models are the following.

Binary Flipping (BF). In this model the bits alternate between the two states: idle and active. At step 0 all of the bits are idling. Let $\chi_1, \chi_2, \dots$ be i.i.d. variables taken from the distribution $\mathcal{P}$. At the ith step, $i = 1, 2, \dots$, the bit with index χ_i is *flipped*, i.e. its state is changed to the opposite:

idle $\leftrightarrow$ *active*

If 0 and 1 represent, respectively, the idling and the active states, the evolution of the system is described by a discrete time Markov chain $\{\zeta_n\}_{n\geq 0} = \{(\zeta_n^1, \zeta_n^2, \zeta_n^3, \dots)\}_{n\geq 0}$ with the state space

$$\mathcal{X} = \left\{x \in \{0,1\}^{\mathbb{N}} : \; x \text{ has finitely many non-zeros}\right\}$$

such that $\zeta_0 = \mathbf{0}$ is the zero-vector, and

$$\zeta_{n+1}^k = \begin{cases} \zeta_n^k, & k \neq \chi_{n+1}, \\ 1 - \zeta_n^k, & k = \chi_{n+1}, \end{cases} \qquad k = 1, 2, \dots, n = 0, 1, 2, \dots . \tag{1}$$

The main quantity of interest is the number of steps required for the system to return to the ground state, i.e. the following stopping time:

$$\begin{aligned} \tau_{\text{BF}} &= \min\{n \geq 1 : \text{no bits are active at step } n\} \\ &= \min\{n \geq 1 : \zeta_n^k = 0, \quad \forall k = 1, 2, \dots\}. \end{aligned}$$

Damaged Bits (DB). This second model elaborates on the first one by adding a damaged state to the bits. As in BF model above, we start with a sequence of idling bits and then consecutively sample from $\mathcal{P}$ for the index of the bit to change its state according to the following dynamics:

idle $\to$ *active* $\to$ *damaged.*

Thus in this model, the reversal of states is not possible: once a bit is active it will never become idle again. An attempt to activate an already active bit leads to its damage. Also, the damaged bits never become functional again and if a damaged bit is selected to change its state nothing happens: it just remains damaged.

If 0,1,2 encode idle, active and damaged states respectively, the corresponding Markov chain $\{\zeta_n\}_{n\geq 0}$ with the state space

$$\mathcal{Y} = \left\{y \in \{0,1,2\}^{\mathbb{N}} : y \text{ has finitely many non-zeros}\right\}$$

is defined by

$$\zeta_{n+1}^k = \begin{cases} \zeta_n^k, & k \neq \chi_{n+1}, \\ \min\{2, 1 + \zeta_n^k\}, & k = \chi_{n+1}, \end{cases} \qquad k, n \in \mathbb{N} \tag{2}$$

with the starting configuration ζ_0 being the vector of all zeroes, $\zeta_0 = \mathbf{0}$.

Here again we are looking for the number of steps to return to the ground state which is now understood as the collection of all of the states without active bits:

$$\begin{aligned} \tau_{\text{DB}} &= \min\{n \geq 1 : \text{no bits are active at step } n\} \\ &= \min\{n \geq 1 : \zeta_n^k \in \{0, 2\}, \quad \forall k = 1, 2, \dots\}. \end{aligned}$$

In contrast to BF model, the ground state in DB model in general cannot be identified with any one particular state of the Markov chain $\{\zeta_n\}$.

Continuous Time Version. So far we have formulated the discrete time dynamics of the system of bits. It is also sensible to consider continuous-time versions of both BF and DB models. Let $\zeta_t = (\zeta_t^1, \zeta_t^2, \dots)$ be a sequence of continuous-time Markov jump processes, each with the state space $\{0, 1\}$ in the BF case, and with $\{0, 1, 2\}$ in the DB case. The kth process $\{\zeta_t^k\}_{t\geq 0}$ represents the corresponding change of states of the kth bit which happens with exponentially distributed holding times at rate p_k. Note that since all p_k sum up to 1, there is an a.s. finite number of state changes of the whole system of bits in any finite period of time. Therefore one can define the renewal process $\{t_n\}$ of times when some of the bits changes its state. The embedded Markov chain $\{\zeta_{t_n}\}_{n\geq 0}$ is then a distributional copy of the discrete-time version $\{\zeta_n\}_{n\geq 0}$ of the model. One of the advantages of this representation, also known as *Poissonisation* and widely used since at least [2], is the independence of ζ_t^k for different $k = 1, 2, \dots$. This often leads to explicitly computable probabilities as we also demonstrate here. Further we use the discrete time and the continuous time versions of the models interchangeably, whichever is more convenient at the moment: the notion of recurrence/transience stays the same for both.

The Markov chains (1) and (2) describing our models can be regarded as random walks on an infinite-dimensional group, see, e.g., [6]. Typically the analysis of random walks on discrete groups assumes a finite generator set, so that the underlying Cayley graph is locally finite, as for example, in [8]. However, the state spaces in our models are not finitely generated groups, so analysis of a random walk in such a space is interesting in its own right. But practical applications are also envisaged: in addition to an evident relation to modelling reliability of a complex system with multiple components prone to fail at different rates, one can also mention computer science and information encryption techniques. The very term "Bit Flipping" is borrowed from the literature on randomised simplex algorithms [3], where a similar model was analysed: each flipped bit there makes all of the bits to the right change their states as well. This model applies to estimate the running speed of a random edge simplex algorithm on a Klee-Minty cube which is particularly 'bad' for many optimisation algorithms thus providing a worst-case scenario, see, e.g., [5] and the references therein.

Finally, we mention and interesting interpretation of the BF model as a dynamical percolation process on $\mathbb{Z}$, where we start with all edges 'open', and then they start 'closing' independently of each other, each with different rate. The question of recurrence is then equivalent to the question of existence of a sequence of percolation times when all the edges are open and thus 0 is connected to the infinity. For a recent survey on the dynamical percolation, see [7].

2 Main Results

For the above models we prove the following main result: each model exhibits a transient or recurrent behaviour, depending on how fast p_k's decay. There is a critical decay separating both regimes, different for each model. We start characterising the critical decay in the Binary Flipping model.

Theorem 2.1. *If the distribution $\mathcal{P}$ is such that:*

(i) $\limsup_{k\to\infty} 2^k p_k < \infty$*, then BF model is recurrent, i.e.* $\mathbf{P}(\tau_{\mathrm{BF}} < \infty) = 1$*,*

(ii) $\liminf_{k\to\infty} (2-\varepsilon)^k p_k > 0$ *for some* $\varepsilon > 0$*, then BF model is transient, i.e.* $\mathbf{P}(\tau_{\mathrm{BF}} = \infty) > 0$*.*

Loosely speaking, the critical decay of p_k's in BF model is the geometric distribution with parameter $\frac{1}{2}$. Although deterministic systems may behave rather differently from stochastic ones, often in a non-critical regime they provide a good intuition to what is happening. Imagine an infinite row of lamps, turning on and off with deterministic frequencies $p_k = (1/2)^{k-1}$. That means, the first lamp changes its state every second, the second lamp every 2 seconds, the third every 4 seconds, etc., meaning that this row is nothing else than a digital clock showing the time since the start in a binary format. Then at least one lamp is lit at every positive time instant. This is still true when $p_k = p^k$ with $p > 1/2$: the $(n+1)$th lamp will always turn on before the nth turns off, so the active intervals of nth and $(n+1)$th lamps will overlap for every n. Thus this deterministic system never returns to the ground state whenever $p \geq 1/2$. However, for $p < 1/2$ the first n lamps will have time to run through all possible combinations (including all zeroes) before the $(n+1)$th lamp will be turned on, so there will always be an infinite number of occurrences of the ground state when no lamp is lit. As Theorem 2.1 shows, the same critical decay separates the stochastic BF model too.

Furthermore, the BF model is never positive recurrent, as the next theorem shows.

Theorem 2.2. *When a BF model is recurrent, it is null-recurrent, i.e.* $\mathbf{E}\,\tau_{\mathrm{BF}} = \infty$ *always.*

This result can be easily foreseen by regarding the BF process as an irreducible time-reversible Markov chain (1). The time-reversibility implies that the stationary measure is uniform, but the state space is countably infinite, hence it cannot be probabilistic so the chain cannot be positive recurrent.

Although the first moment of τ_{BF} is infinite, it is reasonable to ask for which values of $r < 1$ the rth moment becomes finite. The next theorem presents bounds for such r in the case of asymptotically geometrically decaying $\{p_k\}$, these are presented graphically on Figure 1.

Theorem 2.3. *Consider the recurrent BF model in discrete time with $p_k \sim C_1 p^k$ for some fixed constant $C_1 > 0$ and $p \in (0, 1/2)$. Then*

(i) $\mathbf{E}\,\tau_{\mathrm{BF}}^r < \infty$ *for any positive* $r < 1 - \frac{\log 2}{\log(1/p)}$. *Moreover, for any such r, if the Markov chain (1) is started from an arbitrary ζ_0 with the largest active bit M_0, then there exists a constant $C_2 = C_2(C_1, p, r)$ such that*

$$\mathbf{E}\,[\tau_{\mathrm{BF}}^r | M_0 = m] \le C_2 \left(\frac{1}{2p}\right)^m ;$$

(ii) $\mathbf{E}\,\tau_{\mathrm{BF}}^r = \infty$ *for any* $r > 1 - \frac{\log(2-p)}{\log(1/p)}$.

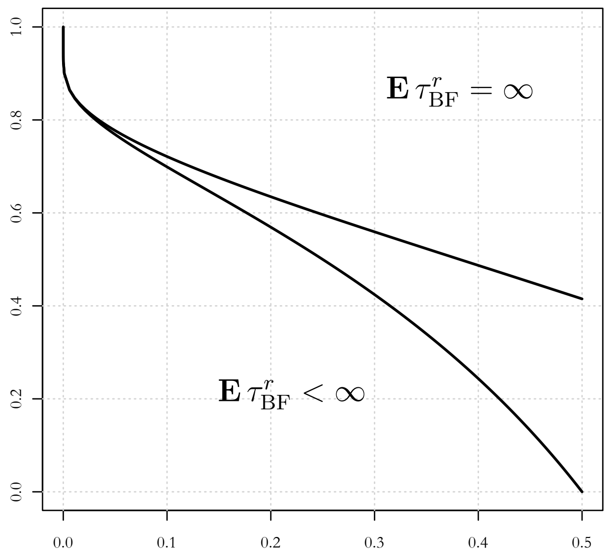

Figure 1: Integrability of τ_{BF}^r as given by Theorem 2.3.

Remark 2.1. *There is an obvious coupling of the DB model with the BF model: just declare the bits which flipped more than once in BF model damaged in DB. Then $\tau_{\mathrm{DB}} \le \tau_{\mathrm{BF}}$ almost surely and the same upper bound* (i) *of Theorem 2.3 is also true for τ_{DB}.*

The DB model can also be recurrent or transient, depending on p_k. The recurrence/transience of the model now does not correspond to recurrence/transience of the Markov chain (2), because the ground state of the DB model is an infinite collection of states of $\{\zeta_n\}$. Still, we call the DB model recurrent, if $\tau_{\mathrm{DB}} < \infty$ with probability 1, and transient otherwise. Denote by Q_k the tail of the distribution $\mathcal{P}$:

$$Q_k = \sum_{j=k+1}^{\infty} p_k.$$

Theorem 2.4. *If the distribution $\mathcal{P}$ is such that:*

(i) $\limsup_{k\to\infty} \frac{Q_{k+1}}{Q_k} = p < 1$, *then the DB model is recurrent,*

(ii) $p_k \sim C\exp(-\alpha k^\gamma), k \to \infty$ *for some* $\alpha > 0, \gamma \in (0, 1/2)$, *then the DB model is transient.*

Denote by η_t the total number of active bits in the continuous version of the model at time $t \geq 0$. In both BF and DB models, whenever $\mathbf{E}\,\eta_t \to \infty$, conditions of the Central Limit Theorem are fulfilled for η_t. We prove the following fact:

Theorem 2.5. *For both BF and DB models, whenever*

$$\mathbf{E}\,\eta_t \to \infty, \tag{3}$$

then also $\mathbf{var}\,\eta_t \to \infty$ *as* $t \to \infty$ *and*

$$\frac{\eta_t - \mathbf{E}\,\eta_t}{\sqrt{\mathbf{var}\,\eta_t}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \quad \text{as } t \to \infty.$$

In BF model the condition (3) *is always fulfilled, and in DB model a sufficient condition for* (3) *is:*

$$p_k \sim C\exp(-\alpha k^\gamma), k \to \infty, \tag{4}$$

for some constants $C > 0, \alpha > 0, \gamma \in (0, 1)$.

Remark 2.2. *In the above theorem, both* $\mathbf{E}\,\eta_t$ *and* $\mathbf{var}\,\eta_t$ *admit an explicit form of a series:*

$$\mathbf{E}\,\eta_t = \sum_{k=1}^{\infty} f(p_k t), \qquad \mathbf{var}\,\eta_t = \sum_{k=1}^{\infty} f(p_k t)(1 - f(p_k t)),$$

where $f(x) = (1 - e^{-x})/2$ *for BF and* $f(x) = xe^{-x}$ *for DB model. In both cases,* $f(p_k t)$ *is the probability for the* k*th bit to be active at time* t *in the corresponding model.*

3 Proofs

3.1 Transience and recurrence of BF model

Proof of Theorem 2.1. First, we are going to prove the theorem for a particular case of $p_k = Cp^k$ for some $p \in (0,1)$ and then extend it using monotonicity arguments.

Consider the continuous-time BF model. Recall $\zeta_t = (\zeta_t^k)_{k\geq 1}$, a continuous-time Markov jump process on $\mathcal{X}$ representing the configuration of the bits at time $t \geq 0$, and $\zeta_0 = \mathbf{0} = (0,0,\dots)$, see (1). Denote by ν_{total} the total time $\{\zeta_t\}$ spends in the state $\mathbf{0}$ for $t > 0$. Since the process $\{\zeta_t\}$ is irreducible, recurrence of the BF model implies that the state $\mathbf{0}$ is recurrent. Since the holding times at state $\mathbf{0}$ are i.i.d. exponential with parameter 1, we get $\mathbf{E}\,\nu_{\text{total}} = \infty$. When the BF model is transient, i.e. when

$$q = \mathbf{P}\{\zeta_t = \mathbf{0} \text{ for some finite } t > t_1 \mid \zeta_0 = \mathbf{0}\} < 1,$$

where t_1 is the time of the first jump of the process ζ_t, then ν_{total} is distributed as the sum $\sum_{i=1}^{\nu} \varepsilon_i$, where ν has geometrical distribution with parameter q and ε_i's are i.i.d. exponentially distributed with parameter 1 r.v.'s representing holding times at state $\mathbf{0}$. In that case, $\mathbf{E}\,\nu_{\text{total}} = \mathbf{E}\,\nu\,\mathbf{E}\,\varepsilon_i = 1/q < \infty$. Thus $\mathbf{E}\,\nu_{\text{total}} = \infty$ is equivalent to recurrence of $\zeta(t)$ and of the BF model.

One can write

$$\mathbf{E}\,\nu_{\text{total}} = \mathbf{E} \int\limits_0^\infty \prod_{k=1}^\infty \mathbb{I}\{k\text{th bit is idle at time } t\}\, dt = \int\limits_0^\infty \prod_{k=1}^\infty \mathbf{P}\{\zeta_t^k = 0\}\, dt.$$

Next,

$$\begin{aligned}\mathbf{P}\{\zeta_t^k = 0\} &= \sum_{j=0}^\infty \mathbf{P}\{k\text{th bit flipped } 2j \text{ times by time } t\}\\ &= e^{-p_k t} \sum_{j=0}^\infty \frac{(p_k t)^{2j}}{(2j)!} = (1 + e^{-2p_k t})/2,\end{aligned}$$

thus the transience is equivalent to the convergence of the integral

$$\mathbf{E}\,\nu_{\text{total}} = \int_0^\infty \prod_{k=1}^\infty (1 + e^{-2p_k t})/2\, dt\,. \tag{5}$$

In the second part of the proof we provide the lower and upper bounds for the infinite product under the integral.

Denote by $f(x) = (1 - e^{-2x})/2$, so that the product under the integral in (5) becomes $\prod\limits_{k=1}^\infty (1 - f(p_k t))$.

Fix an arbitrary small $\varepsilon > 0$. Note that the function $1 - f(x)$ is monotone decreasing in x and the equation

$$1 - f(x) = \frac{1}{2 - \varepsilon}$$

has the only root. Call this root z_ε. Now, represent the product as a multiplication of the two factors:

$$\prod_{k=1}^{\infty}(1 - f(p_k t)) = \underbrace{\prod_{k:p_k t < z_\varepsilon}(1 - f(p_k t))}_{\Phi_1(t)} \underbrace{\prod_{k:p_k t \geq z_\varepsilon}(1 - f(p_k t))}_{\Phi_2(t)},$$

First, note that every term of the product $\Phi_1(t)$ is less or equal than 1, therefore, $\Phi_1(t) \leq 1$. Next, observe that $\{p_k\}_{k\geq 1}$ is the geometric distribution, and therefore $\Phi_1(t) = \Phi_1(\frac{t}{p^n})$ for each $n = 1, 2, \ldots$, moreover, taking into account that the function f is continuous, non-increasing, we obtain

$$\begin{aligned}
\Phi_1(t) &= \prod_{k:p_k t < z_\varepsilon}(1 - f(p_k t)) \geq \prod_{k=1}^{\infty}(1 - f(z_\varepsilon p_k)) \\
&= \exp\left\{\sum_{k=1}^{\infty}\log(1 - f(p^k z_\varepsilon))\right\} \geq \exp\left\{-\sum_{k=1}^{\infty} f(p^k z_\varepsilon)\right\} \\
&= \exp\left\{-\frac{1}{2}\sum_{k=1}^{\infty}(1 - e^{-2p^k z_\varepsilon})\right\} \geq \exp\left\{-\frac{1}{2}\sum_{k=1}^{\infty} 2p^k z_\varepsilon\right\} \\
&= \exp\left\{-\frac{z_\varepsilon}{1 - p}\right\}
\end{aligned}$$

Therefore for any positive t, $C_1 < \Phi_1(t) < C_2$ with fixed and finite positive constants C_1, C_2.

As for the second factor $\Phi_2(t)$, if we denote $A(t) = \{k : p_k t \geq z_\varepsilon\}$, then for any $k \in A(t)$ we have $1 - f(p_k t) \leq 1/(2 - \varepsilon)$, and thus

$$\left(\frac{1}{2}\right)^{|A(t)|} \leq \Phi_2(t) \leq \left(\frac{1}{2 - \varepsilon}\right)^{|A(t)|}.$$

Since $|A(t)| = \mathrm{card}\{k : p_k \geq \frac{z_\varepsilon}{t}\} = \mathrm{card}\{k : k < \frac{\log z_\varepsilon}{\log p} - \frac{\log Ct}{\log p}\} = C_3 + \lfloor \frac{\log t}{\log \frac{1}{p}} \rfloor$, we obtain

$$C_4\left(\frac{1}{2}\right)^{\frac{\log t}{\log \frac{1}{p}}} < \Phi_2(t) < C_5\left(\frac{1}{2 - \varepsilon}\right)^{\frac{\log t}{\log \frac{1}{p}}},$$

and finally,

$$C_6 t^{-\frac{\log 2}{\log \frac{1}{p}}} < \prod_{k=1}^{\infty}(1 - f(p_k t)) < C_7 t^{-\frac{\log(2-\varepsilon)}{\log \frac{1}{p}}},$$

which yields the theorem statement for geometric $\{p_k\}$, recalling an arbitrary small choice of ε.

Moving to a general $\{p_k\}$, in case (i) for all sufficiently large k, $p_k < C_8 2^{-k} < 2^{C_9 - k}$, and since $1 - f(x)$ is non-increasing in x, and $1 - f(x) > 1/2$ for $x > 0$, we can choose a large enough M and write

$$\begin{aligned}
\int_0^\infty \prod_{k=1}^\infty (1 - f(p_k t))\, dt &\geq C_{10} \int_0^\infty \prod_{k=M}^\infty (1 - f(p_k t))\, dt \\
&\geq C_{10} \int_0^\infty (1 - f(2^{C_9 - k} t))\, dt \\
&= C_{10} \int_0^\infty \prod_{k=1}^\infty (1 - f(2^{-k} \cdot 2^{C_9 + M - 1} t))\, \frac{d(2^{C_9+M-1} t)}{2^{C_9+M-1}} \\
&= C_{11} \int_0^\infty \prod_{k=1}^\infty (1 - f(2^{-k} t))\, dt.
\end{aligned}$$

Similarly, in case (ii), for all sufficiently large k, $p_k > C_{12}(2 - \varepsilon)^{-k} > (2 - \varepsilon)^{C_{13} - k}$, and $1 - f(x) \leq 1, x > 0$, so we can choose a sufficiently large M so that

$$\int_0^\infty \prod_{k=1}^\infty (1 - f(p_k t))\, dt \leq C_{14} \int_0^\infty \prod_{k=1}^\infty (1 - f((2 - \varepsilon)^{-k} t))\, dt,$$

and the theorem statement follows. $\square$

We have seen that a BF model can be recurrent, but can it be positive recurrent, i.e. can the number of steps to return to the ground state have a finite expectation? The negative answer is provided by Theorem 2.2 which we prove next.

Proof of Theorem 2.2.

Introduce the following notation:

$$\zeta_n^{\wedge m} = (\zeta_n^1, \ldots, \zeta_n^m), \quad \mathbf{0}^{\wedge m} = (\underbrace{0, \ldots, 0}_{m}), \; \tau_{\mathrm{BF}}^{\wedge m} = \inf\{n \in \mathbb{N} : \zeta_n^{\wedge m} = \mathbf{0}^{\wedge m}\}$$

Obviously, $\tau_{\mathrm{BF}} = \inf\{n \in \mathbb{N} : \zeta_n = \mathbf{0}\} \geq \tau_{\mathrm{BF}}^{\wedge m}$ almost surely. Next, $\{\zeta_n^{\wedge m}\}_{n \geq 0}$ is an irreducible aperiodic (contrary to $\{\zeta_n\}_{n \geq 0}$, which has period 2) Markov chain with the finite state space $\{0, 1\}^m$ and a symmetric transition matrix, hence its unique stationary distribution $\pi^{\wedge m}$, given by the solution of the detailed balance equations

$$\pi^{\wedge m}(x) p(x, y) = \pi^{\wedge m}(y) p(y, x), \quad x, y \in \{0, 1\}^m,$$

is uniform on $\{0, 1\}^m$. Consequently, $\pi^{\wedge m}(\mathbf{0}^{\wedge m}) = 2^{-m}$ and $\mathbf{E}\, \tau_{\mathrm{BF}}^{\wedge m} = (\pi^{\wedge m}(\mathbf{0}^{\wedge m}))^{-1} = 2^m$.

Finally,

$$\mathbf{E}\,\tau_{\mathrm{BF}} \geq \mathbf{E}\,\tau_{\mathrm{BF}}^{\wedge m} = 2^m \text{ for every } m \in \mathbb{N},$$

finishing the proof. $\square$

In order to prove Theorem 2.3, we make use of Theorem 1 and Corollary 1 in [1]. For convenience of the reader, we give their formulation in our notation.

Theorem 3.1 ([1, Theorem 1])**.** *Suppose that $\{Y_n\}_{n\geq 0}$ is an $\{\mathcal{F}_n\}$-adapted stochastic process taking values in an unbounded subset of $\mathbb{R}_+$. Introduce $\tau_A = \inf\{n \geq 0 : Y_n \leq A\}$. Suppose there exist positive constants A, ε such that for every n, Y_n^{2r} is integrable and*

$$Y_n^{2-2r}\mathbf{E}\left[Y_{n+1}^{2r} - Y_n^{2r} \,\middle|\, \mathcal{F}_n\right] \leq -\varepsilon \quad \text{on } \{\tau_A \geq n\}. \tag{6}$$

Then for any $r*$ satisfying $0 < r^* < r$ there exists a constant $c = c(\varepsilon, r^*, r)$ such that for any $x \geq 0$

$$\mathbf{E}\,\tau_A^{r^*} \leq cx^{2r} \quad \text{whenever } Y_0 = x \text{ a.s.}$$

Theorem 3.2 ([1, Corollary 1])**.** *Let $\{Y_n\}_{n\geq 0}$, τ_A be as in Theorem 3.1. Suppose there exist positive constants A, ε, and J such that for any n,*

$$\mathbf{E}\,[Y_{n+1}^2 - Y_n^2 \mid \mathcal{F}_n] \geq -\varepsilon \text{ on } \{\tau_A > n\}$$

and, for some $\rho > 1$,

$$Y_n^{2-2\rho}\mathbf{E}\,[Y_{n+1}^{2\rho} - Y_n^{2\rho} \mid \mathcal{F}_n] \leq J \text{ on } \{\tau_A > n\}.$$

Suppose also that $Y_0 = x > A$ and for some positive r_0 the process $\{Y_{n\wedge\tau_A}^{2r_0}\}_{n\geq 0}$ is a submartingale. Then for any $r > r_0$, $\mathbf{E}\,\tau_A^r = \infty$.

We will also need the following technical Lemma.

Lemma 3.1. *Let $\{\zeta_n\}_{n\geq 0}$ be a discrete time BF model starting from the ground state $\zeta_0 = \mathbf{0}$ with the parameter distribution $\mathcal{P} = \{p_1, p_2, \dots\}$ possibly with a finite support: $p_1 \geq p_2 \geq p_3 \geq \ldots \geq 0$. Then for $K = \min\{k : \sum_{i=k}^{\infty} p_i \leq 1/2\}$ and any $n = 1, 2, \dots$, the vector $(\zeta_n^K, \zeta_n^{K+1}, \dots)$ is stochastically dominated by the vector $(\check{\zeta}^K, \check{\zeta}^{K+1}, \dots)$ of i.i.d.* Bern(1/2) *random variables.*

Proof of Lemma 3.1. Assume that $\sum_{k=K}^{\infty} p_k > 0$, otherwise the Lemma statement is trivial. Let $\{\zeta_n\}_{n\geq 0}, \{\check{\zeta}_n\}_{n\geq 0}$ be two discrete time BF models with the same transition probabilities

$$\mathbf{P}\{\zeta_{n+1}^k = 1 - \zeta_n^k\} = \mathbf{P}\{\check{\zeta}_{n+1}^k = 1 - \check{\zeta}_n^k\} = p_k, \; k \in \mathbb{N}, \; n = 0, 1, \dots, \tag{7}$$

where the first one starts from the ground state $\zeta_0 = \mathbf{0}$ and the second one starts from the stationarity: $\check{\zeta}_0$ is a sequence of i.i.d. symmetric Bernoulli random variables:

$$\mathbf{P}\{\check{\zeta}_0^k = 1\} = 1 - \mathbf{P}\{\check{\zeta}_0^k = 0\} = 1/2 \quad \text{for all } k \in \mathbb{N}.$$

Obviously, $0 = \zeta_0^k \le \check{\zeta}_0^k$ for all $k \in \mathbb{N}$ almost surely. Our goal is to couple the Markov chains $\{\zeta_n\}$ and $\{\check{\zeta}_n\}$ on $\{0,1\}^{\mathbb{N}}$ preserving the almost sure coordinate-wise domination $\zeta_n^k \le \check{\zeta}_n^k$ for all n and all $k = K, K+1, \ldots$.

The idea is to treat the first $(K-1)$ bits of both Markov chains as a kind of 'buffer' for which the domination does not generally holds. This is an expense to pay for the donination for the large coordinates.

Specifically, we define the joint transition dynamics for $\zeta_n, \check{\zeta}_n$ inductively, for $n = 0, 1, 2, \ldots$. Denote by D_n the (random) set of discrepancies at time n, i.e. the set of indices $k \ge K$ at which $\zeta_n, \check{\zeta}_n$ disagree. The induction assumption is that the coordinate-wise domination is preserved on step n: $\zeta_n^k \le \check{\zeta}_n^k$ for all $k \ge K$ and hence only discrepancies of the form $\zeta_n^k = 0$, $\check{\zeta}_n^k = 1$ are possible. Denote these by

$$D_n = \{k \ge K : \zeta_n^k = 0,\ \check{\zeta}_n^k = 1\}.$$

The domination obviously holds for $n = 0$.

Let $F^{-1}(u)$ be the quantile function for the distribution $\mathcal{P}$:

$$F^{-1}(u) = \min\left\{k : \sum_{i=1}^{k} p_i > u\right\}, \quad u \in (0,1).$$

The key element of the construction is a map $s_n(u) : (0,1) \to (0,1)$ which swaps the parts of $(0,1)$ corresponding to D_n with the parts of $(0,1)$ of the same length, corresponding to the buffer:

$$s_n(u) = \begin{cases} 1-u, & \text{if } F^{-1}(u) \in D_n, \text{ or } F^{-1}(1-u) \in D_n \\ u, & \text{otherwise.} \end{cases}$$

Introduce a common source of randomness for the chains: the sequence $U_1, U_2, \ldots$ of i.i.d. random variables distributed uniformly on the interval $(0,1)$. The indices of the bits to flip on step $(n+1)$ in ζ_n and $\check{\zeta}_n$, $n = 0, 1, \ldots$, are defined, respectively, as

$$\begin{aligned} \chi_{n+1} &= F^{-1}(U_{n+1}), \\ \check{\chi}_{n+1} &= F^{-1}(s_n(U_{n+1})). \end{aligned}$$

Since $s_n(u)$ preserves the Lebesgue measure, $s_n(U_{n+1})$ is also uniformly distributed implying that both chains have correct transition probabilities:

$$\mathbf{P}(\chi_{n+1} = k) = \mathbf{P}(\check{\chi}_{n+1} = k) = p_k, k = 1, 2, \ldots, n = 0, 1, 2, \ldots$$

Moreover, if one of the chains is flipped at some coordinate $k \geq K$, where the chains agree, the other one does the same. If, otherwise, one of the chains is selected to be flipped at some coordinate $k \geq K$, where the chains disagree, the other one is flipped at one of the coordinates $k = 1, 2, \ldots, K-1$ of the buffer. As a result, the chains will agree at the flipped coordinate from D_n. Thus no new discrepancies are created for $k \geq K$ and the coordinate-wise domination $\zeta_{n+1}^k \leq \check{\zeta}_{n+1}^k$ is preserved almost surely.

$\square$

Proof of Theorem 2.3. Part (i). Denote by M_n the index of the rightmost active bit at time n: $M_n = \max\{k : \zeta_n^k = 1\}$ with convention $M_n = 0$ for $\zeta_n = \mathbf{0}$. Put $Y_n^{2r} = y^{M_n}$ for some $y > 1$ which will be selected later. Define the filtration $\mathcal{F}_n = \sigma(\zeta_0, M_1, \ldots, M_n)$. The process $\{Y_n\}$ is obviously adapted to $\{\mathcal{F}_n\}$. Recall that χ_k is an index of a bit flipped on step k, $\chi_k \sim \mathcal{P}$ by the assumptions of (i). We have that

$$\mathbf{E}\left(Y_n^{2r}\right) = \mathbf{E}\left(y^{M_n}\right) \leq \mathbf{E}\left(y^{\sum_{k=1}^n \chi_k}\right) = \left(\mathbf{E}\left(y^{\chi_1}\right)\right)^n. \tag{8}$$

The inequality above follows, since $M_n \leq \max\{\chi_1, \chi_2, \ldots, \chi_n\} \leq \sum_{k=1}^n \chi_k$, so that the right-hand side of (8) is finite whenever

$$py < 1. \tag{9}$$

Next,

$$\begin{aligned}\mathbf{E}\left[Y_{n+1}^{2r} - Y_n^{2r} \;\middle|\; M_n = m\right] &= \underbrace{\mathbf{E}\left[\left(Y_{n+1}^{2r} - Y_n^{2r}\right) \mathbb{I}_{\{\chi_{n+1} = m\}} \;\middle|\; M_n = m\right]}_{E_1} \\ &\quad + \underbrace{\mathbf{E}\left[\left(Y_{n+1}^{2r} - Y_n^{2r}\right) \mathbb{I}_{\{\chi_{n+1} > m\}} \;\middle|\; M_n = m\right]}_{E_2}.\end{aligned}$$

Introduce $\psi(x_K, \ldots, x_{m-1}) = y^{\max\{j:\; x_j = 1,\; j = K, \ldots, m-1\}} - y^m$. Then

$$\begin{aligned}E_1 &\leq \mathbf{E}\left[\psi(\zeta_n^K, \ldots, \zeta_n^{m-1})\, \mathbb{I}\{\chi_{n+1} = m\} \;\middle|\; M_n = m\right] \\ &= p_m \mathbf{E}\left[\psi(\zeta_n^K, \ldots, \zeta_n^{m-1}) \;\middle|\; M_n = m\right]\end{aligned} \tag{10}$$

Our next step is to show that the vector $(\zeta_n^K, \ldots, \zeta_n^{m-1})$ conditionally on $\{M_n = m\}$ is stochastically dominated by a vector of i.i.d. Bernoulli random variables $(\check{\zeta}^K, \ldots, \check{\zeta}^{m-1})$. Introduce an embedded Markov chain

$$\left\{\widetilde{\zeta}_l\right\}_{l \geq 0} = \left\{(\widetilde{\zeta}_l^1, \ldots, \widetilde{\zeta}_l^{m-1})\right\}_{l \geq 0}$$

tracking the state of the first $(m-1)$ coordinates of $\{\zeta_n\}$ considered at the times when one of those coordinates changes. We set $\widetilde{\zeta}_0 = (\zeta_0^1, \ldots, \zeta_0^{m-1})$

and define $\widetilde{\zeta}_l = (\zeta^1_{t_l(m)}, \ldots, \zeta^{m-1}_{t_l(m)})$, where $t_l(m)$ is the lth time when one of the first $(m-1)$ coordinates of ζ_n is flipped.

Lemma 3.1 applied to the BF model $\{\widetilde{\zeta}_l\}_{l\geq 0}$ with the corresponding flipping probabilities

$$\widetilde{\mathcal{P}} = \left\{ \frac{p_1}{S_{m-1}}, \ldots, \frac{p_{m-1}}{S_{m-1}}, 0, 0, \ldots \right\}, \quad S_{m-1} = \sum_{k=1}^{m-1} p_k,$$

implies for every $l = 0, 1, 2, \ldots$ the stochastic domination

$$(\widetilde{\zeta}_l^{\tilde{K}}, \ldots, \widetilde{\zeta}_l^{m-1}) \leq_{\mathrm{st}} (\check{\zeta}^{\tilde{K}}, \ldots, \check{\zeta}^{m-1}),$$

where $\check{\zeta}^{\tilde{K}}, \ldots, \check{\zeta}^{m-1}$ are i.i.d. $\mathsf{Bern}(1/2)$ random variables. Note that

$$\tilde{K} = \min\{k : \sum_{i=k}^{\infty} p_i \leq 1/2\, S_{m-1}\} \leq K = \min\{k : \sum_{i=k}^{\infty} p_i \leq 1/2\}.$$

Therefore, for every $l = 0, 1, 2, \ldots,$

$$(\widetilde{\zeta}_l^{K}, \ldots, \widetilde{\zeta}_l^{m-1}) \leq_{\mathrm{st}} (\check{\zeta}^{K}, \ldots, \check{\zeta}^{m-1}). \tag{11}$$

Introduce the series of events:

$$\begin{aligned} A(n,m,l) =& \{\text{by the time } n \text{ the first } m-1 \text{ coordinates of } \zeta \\ & \text{are flipped } l \text{ times}\} \\ =& \Big\{\sum_{k=1}^{n} \mathbb{I}\{1 \leq \chi_k \leq m-1\} = l\Big\}, \end{aligned}$$

for $n = 0, 1 \ldots,$ and $l = 0, \ldots, n$. Conditionally on $A(n,m,l)$, the distribution of $(\zeta^1_n, \ldots, \zeta^{m-1}_n)$ is the same as that of $(\widetilde{\zeta}^1_l, \ldots, \widetilde{\zeta}^{m-1}_l)$, so we can continue (10) with:

$$\begin{aligned} E_1 &\leq p_m \sum_{l=0}^{n} \mathbf{E}\left[\psi(\zeta_n^K, \ldots, \zeta_n^{m-1})\, \mathbb{I}_{A(n,m,l)} \;\Big|\; M_n = m\right] \\ &= p_m \sum_{l=0}^{n} \mathbf{E}\left[\psi(\zeta_n^K, \ldots, \zeta_n^{m-1})\, \mathbb{I}_{M_n = m} \;\Big|\; A(n,m,l)\right] \frac{\mathbf{P}\,(A(n,m,l))}{\mathbf{P}(M_n = m)}. \end{aligned}$$

Now notice, that conditionally on $A(n,m,l)$, random variables $\psi(\zeta_n^K, \ldots, \zeta_n^{m-1}) = \psi(\widetilde{\zeta}_l^K, \ldots, \widetilde{\zeta}_l^{m-1})$ and $\mathbb{I}_{M_n=m}$ are independent. Indeed, on $A(n,m,l)$, the first variable is a function of the chain $\widetilde{\zeta}$ after l steps which is governed by transition probabilities $\widetilde{\mathcal{P}}$. While the event $M_n = m$ relates to configuration of the bits $m, m+1, \ldots$ after $n-l$

steps of the BF model with parameter distribution $\{p_k/(1-S_{m-1}),\ k = m, m+1, \dots\}$. Therefore,

$$\begin{aligned}E_1 \le & p_m \sum_{l=0}^{n} \mathbf{E}\left[\psi(\zeta_n^K, \dots, \zeta_n^{m-1}) \bigm| A(n,m,l)\right] \times \\ & \mathbf{P}\bigl(M_n = m \bigm| A(n,m,l)\bigr) \frac{\mathbf{P}\bigl(A(n,m,l)\bigr)}{\mathbf{P}(M_n = m)} \\ = & p_m \sum_{l=0}^{n} \mathbf{E}\left[\psi(\widetilde{\zeta}_l^K, \dots, \widetilde{\zeta}_l^{m-1})\right] \mathbf{P}\bigl(A(n,m,l) \bigm| M_n = m\bigr).\end{aligned}$$

The function ψ is non-decreasing with respect to the coordinate-wise order on its argument, so the stochastic domination (11) implies:

$$\begin{aligned}E_1 &\le p_m \mathbf{E}\, \psi(\check{\zeta}^K, \dots, \check{\zeta}^{m-1}) \underbrace{\sum_{l=0}^{n} \mathbf{P}\left(A(n,m,l) \bigm| M_n = m\right)}_{=1} \\ &= p_m \mathbf{E}\, \psi(\check{\zeta}^K, \dots, \check{\zeta}^{m-1}) \\ &= \sum_{k=K}^{m-1} (y^k - y^m) p_m \left(\frac{1}{2}\right)^{m-k}.\end{aligned}$$

Fix an arbitrary small $\varepsilon > 0$. Since $p_k \sim C_1 p^k, k \to \infty$, one can, if necessary, increase K so that $p_k \ge C_1(1-\varepsilon)p^k$ for any $k \ge K$, and continue:

$$\begin{aligned}E_1 &\le C_1(1-\varepsilon)(py)^m \sum_{k=K}^{m-1} ((2y)^{-k} - 2^{-m+k}) \\ &= C_1(1-\varepsilon)(py)^m \left(\frac{2-2y}{2y-1} - \frac{(2y)^{-m+K}}{2y-1} + 2^{-m+K}\right) \\ &\le C_1(1-\varepsilon)(py)^m \left(\frac{2-2y}{2y-1} + 2^{-m+K}\right).\end{aligned}$$

Because of our assumption $py < 1$, and the asymptotic equivalence $p_k \sim C_1 p^k, k \to \infty$, for an arbitrary small $\varepsilon > 0$ we can choose a large enough $M = M(\varepsilon)$ so that for any $m \ge M$:

$$\begin{aligned}E_2 &= \sum_{k=1}^{\infty} p_{m+k}(y^{m+k} - y^m) \le C_1(1+\varepsilon)(py)^m \left(\sum_{k=1}^{\infty}(py)^k - \sum_{k=1}^{\infty} p^k\right) \\ &= C_1(1+\varepsilon)(py)^m \left(\frac{py}{1-py} - \frac{p}{1-p}\right).\end{aligned}$$

Introduce

$$Q(p,y,\varepsilon) = (1-\varepsilon)\frac{2-2y}{2y-1} + (1+\varepsilon)\Bigl(\frac{py}{1-py} - \frac{p}{1-p}\Bigr).$$

Then $Y_n^{2r} = y^{M_n}$ yields

$$Y_n^{2-2r}\mathbf{E}\left[Y_{n+1}^{2r} - Y_n^{2r} \;\middle|\; M_n = m\right] \le C_1(Q(p,y,\varepsilon)+(1-\varepsilon)2^{-m+2})(py^{\frac{1}{r}})^m.$$

Now fix a $p < 1/2$. For the last expression to be negative and separated from zero for all m large enough, it is necessary for $Q(p,y,\varepsilon)$ to be negative and for $py^{\frac{1}{r}}$ to be greater than one. However, $Q(p,y,\varepsilon) < 0$ reduces to

$$\begin{cases} 0 < p < 1/2, \\ 1 < y < \frac{1}{2p}, \\ 0 < \varepsilon < \frac{2p^2y-4py-p+2}{2p^2y-3p+2}. \end{cases}$$

The right part of the third inequality is positive whenever the first two inequalities are satisfied. Putting it all together, for a fixed pair of p, r we can pick y and M so that $Y_n = y^{\frac{M_n}{2r}}$, given that $M_0 > M$, satisfies the conditions of Theorem 3.1 if and only if the following system of inequalities can be solved for y:

$$\begin{cases} 1 < y < \frac{1}{2p}, \\ py^{\frac{1}{r}} > 1, \end{cases}$$

and the latter is possible when $r < 1 - \frac{\log 2}{\log \frac{1}{p}}$.

Denote $\tau_x = \inf\{n \ge 1 : M_n \le x\}$. Then Theorem 3.1 implies that for $p < 1/2$ and $r < 1 - \frac{\log 2}{\log \frac{1}{p}}$ there exists $C = C(p,r)$ such that for a particular choice of y, M we have

$$\mathbf{E}\left[\tau_M^r \;\middle|\; M_0 = x\right] \le Cy^x \le C\left(\frac{1}{2p}\right)^x.$$

Now we prove that $\tau_{\mathrm{BF}}^r = \tau_0^r$ is integrable and satisfies the same asymptotic bound. In $\mathbf{E}[\tau_0^r | M_0 = x]$, τ_0 is the first time when the process M_n reaches 0 starting from the state x. For any $M \ge 0$ we have $\tau_0 = \tau_M + (\tau_M - \tau_0)$. By simple coupling arguments, the law of $(\tau_M - \tau_0)$, conditional on $\{M_0 = x\}$, is stochastically dominated by the law of τ_0, conditional on $\{M_0 = M\}$. That, together with the inequality $(a+b)^r \le 2^r(a^r + b^r)$ for $0 < r < 1$ and non-negative a, b, gives the bound:

$$\mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = x\right] \le 2^r\big(\mathbf{E}\left[\tau_M^r \;\middle|\; M_0 = x\right] + \mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = M\right]\big).$$

We have just obtained an asymptotic upper bound for the first conditional expectation under the parentheses. It is left now to show that the second expectation is also bounded. Conditionally on $\{M_0 = M\}$, τ_0 is stochastically dominated from above by the sum of two terms. The first one is the time needed for ζ_n to reach $\mathbf{0}$ not leaving the finite sub-cube $\{0,1\}^M$, which is in turn dominated by $\tau_0^{\wedge M} = \inf\{n : \zeta_n^{\wedge M} = \mathbf{0}^{\wedge M}\}$. The second one is a geometrically distributed number of excursions $\gamma \sim \mathrm{Geom}(\pi)$ from $\{0,1\}^M$. Thus

$$\mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = M\right] \le \sum_{k=1}^{\infty} \mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = M, \gamma = k\right] \mathbf{P}\{\gamma = k\}.$$

Now, conditionally on $\{\gamma = k\}$,

$$\begin{aligned}\mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = M, \gamma = k\right] &\le \mathbf{E}\left[\left(\tau_0^{\wedge M} + \sum_{j=1}^{k} \psi_j\right)^r \;\middle|\; M_0 = M\right] \\ &\le k^{1+r}\left(\mathbf{E}\left[(\tau_0^{\wedge M})^r \;\middle|\; M_0 = M\right] + \mathbf{E}\,\psi^r\right),\end{aligned}$$

where ψ_j is the length of excursion $j = 1, \dots, \gamma$ and ψ stands for length of a typical excursion. The first expectation inside the parentheses is a finite constant. As for the second, for some constant $C_2 > 0$ we have

$$\begin{aligned}\mathbf{E}\,\psi^r &= 1 + \sum_{k=1}^{\infty} p_{k+M} \mathbf{E}\left[\tau_M^r \;\middle|\; M_0 = k + M\right] \\ &\le 1 + \sum_{k=1}^{\infty} C_2 p^{k+M} \left(\frac{1}{2p}\right)^{k+M} < \infty.\end{aligned}$$

Thus for some constant $C_3 > 0$

$$\mathbf{E}\left[\tau_0^r \;\middle|\; M_0 = M\right] \le \sum_{k=1}^{\infty} C_3 k^{1+r} \pi (1-\pi)^{k-1} < \infty,$$

finishing the proof of part (i).

Proof of part (ii). Put $Y_n^2 = y^{M_n}$ for some $y > 1$ and check the conditions of Theorem 3.2. As before, Y_n is adapted and for an arbitrary small $\varepsilon > 0$ we can choose $M = M(\varepsilon)$ large enough so that:

$$\begin{aligned}\mathbf{E}\left[Y_{n+1}^2 - Y_n^2 \;\middle|\; M_n = m\right] \ge &- p_m y^m + \sum_{k=1}^{\infty} p_{m+k}(y^{m+k} - y^m) \\ = &- C_1(1-\varepsilon)p^m y^m \\ &+ \sum_{k=1}^{\infty} C_1(1+\varepsilon)p^{m+k}(y^{m+k} - y^m) \\ =&C_1(py)^m\left(-1 + \varepsilon + (1+\varepsilon)\sum_{k=1}^{\infty} p^k(y^k - 1)\right) \\ =&C_1(py)^m\left(-1 + \varepsilon + \frac{(1+\varepsilon)p(-1+y)}{(1-p)(1-py)}\right) \\ =&C_1(py)^m R(p, y, \varepsilon),\end{aligned}$$

where $R(p,y,\varepsilon) = \left(-1+\varepsilon+\frac{(1+\varepsilon)p(-1+y)}{(1-p)(1-py)}\right)$. It is then possible to choose a small enough $\varepsilon > 0$ and a large M so that the latter expression is bounded from below for all $m > M$, when $py < 1$. Furthermore, for such p, y we have as before:

$$Y_n^{2-2\rho}\mathbf{E}\left[Y_{n+1}^{2\rho} - Y_n^{2\rho} \;\middle|\; M_n = m\right] \le C_1 y^{m(1-\rho)}(py^\rho)^m \times \\ (Q(p, y^\rho, \varepsilon) + (1-\varepsilon)2^{-m+K})$$

which is bounded from above when ρ is such that $py^\rho < 1$ (such a $\rho > 1$ exists whenever $py < 1$).

Finally, check for which r_0 the process $Y_{n\wedge\tau_M}^{2r_0}$ is a submartingale. Since

$$\mathbf{E}\left[Y_{n+1}^{2r_0} - Y_n^{2r_0} \;\middle|\; M_n = m\right] \ge C_1(py^{r_0})^m R(p, y^{r_0}, \varepsilon),$$

we can choose $\varepsilon > 0$ so that the latter is greater than zero for any $m > M$, if $r_0 \in (\frac{\log\frac{1}{2p-p^2}}{\log y}, 1)$. Recalling that we can take y arbitrary close to $1/p$, we conclude that the conditions of Theorem 3.2 are satisfied for any r_0 such that $r_0 \in (1 - \frac{\log(2-p)}{\log\frac{1}{p}}, 1)$. Adding this together with the results of Theorem 2.2 implies that none of the fractional moments of τ_M (and hence of τ_0) of order higher than $1 - \frac{\log(2-p)}{\log\frac{1}{p}}$ exists, finishing the proof of Part (ii). □

3.2 Transience and recurrence of DB model

Proof of Theorem 2.4. (i) Consider the discrete-time version of the DB model. Introduce R_n – the index of the rightmost bit (i.e. with the largest index) that has ever been flipped by time n. The sequence $\{R_n\}$ is a.s. non-decreasing. We aim to prove that almost surely for infinitely many terms of the sequence $\{R_n\}$, each of the bits $1, 2, \ldots, R_n$ is flipped at least twice before the next flip of some bit with an index larger than R_n. That would guarantee that the ground state of the DB model, corresponding to the set of states

$$\left\{y \in \{0,1,2\}^{\mathbb{N}} : y \text{ has no 1's and only a finite number of 2's}\right\}$$

of Markov chain $\{\zeta_n\}$, is visited infinitely often.

It is convenient to use the continuous-time representation now. Let $\Pi_1(t), \Pi_2(t), \ldots$ be the sequence of independent Poisson processes (clocks) describing the times at which, respectively, the 1st, the 2nd, etc. bits are flipped. Introduce $\tau_{>k} = \inf\{t > 0 : \sum_{j=k+1}^{\infty}\Pi_j(t) > 0\}$, the time of the first flip of a bit with an index greater than k. Note that $\tau_{>k}$ is a stopping time for each $k = 0, 1, 2, \ldots$, and, moreover,

$$\tau_{>1} \le \tau_{>2} \le \tau_{>3} \le \ldots$$

Introduce the events

$$\begin{aligned}
A_k &= \quad \{k\text{th bit appears in the sequence } \{R_n\}\}, \\
B_k &= \quad A_k \cap \{\text{starting from the first flip of } k\text{'th bit, each of the bits} \\
&\qquad\quad 1, 2, \ldots, k \text{ is flipped at least twice before the first flip} \\
&\qquad\quad \text{of one of the bits } k+1, k+2, \ldots\}
\end{aligned}$$

Our aim is to prove that the events B_k happen infinitely often. In terms of a continuous-time notation, we can rewrite:

$$\begin{aligned}
A_k &= \{\tau_{>k-1} < \tau_{>k}\}, \\
B_k &= \bigcap_{j \le k} \{\Pi_j(\tau_{>k-1}, \tau_{>k}) \ge 2\},
\end{aligned} \tag{12}$$

where $\Pi(t_1, t_2)$ stands for the number of points a Poisson process Π has in (t_1, t_2). Since $\{\tau_{>k}\}$ is a sequence of stopping times, it is not hard now to see that the events B_k are independent of each other. By the Borel–Cantelli Lemma it suffices to prove that the series $\sum_{k \ge 1} \mathbf{P}\{B_k\}$ diverges.

The probability of A_k (probability of an index k to ever appear in the sequence $\{R_n\}$) is exactly $p_k/(p_k + Q_k) = 1 - Q_k/Q_{k-1}$, which is uniformly bounded away from zero given assumptions of (i).

As follows from (12), the probability $\mathbf{P}(B_k \mid A_k)$ is equal to the probability for each of the first k Poisson clocks $\Pi_1(t), \ldots, \Pi_k(t)$ to tick at least twice before the time of the first tick of one of the clocks $\Pi_{k+1}(t), \Pi_{k+2}(t), \ldots$ We write:

$$\begin{aligned}
\mathbf{P}(B_k \mid A_k) &= \mathbf{P}\Big(\bigcap_{j=1}^{k} \{\Pi_j(\tau_{>k}) \ge 2\}\Big) \\
&= \int_0^\infty \prod_{j=1}^{k} \mathbf{P}\{\Pi_j(t) \ge 2\} \, d\mathbf{P}(\tau_{>k} \le t)
\end{aligned} \tag{13}$$

Introduce $g(x) = e^{-x}(1+x)$. Now, due to (i), there exists a large K such that for any $k \ge j \ge K$:

$$\frac{p_j}{Q_k} = \frac{p_j}{Q_{j-1}} \frac{Q_{j-1}}{Q_j} \ldots \frac{Q_{k-1}}{Q_k} \ge \left(\frac{1}{p} - 1\right) \underbrace{\cdot \frac{1}{p} \ldots \frac{1}{p}}_{k-j+1} = C_2 p^{j-k}.$$

The function $g(x)$ is strictly decreasing in x, so we can continue (13) :

$$
\begin{aligned}
&= \int_0^\infty \prod_{j=1}^{k} (1 - g(p_j t)) Q_k e^{-Q_k t} \, dt \\
&= \int_0^\infty \prod_{j=1}^{k} (1 - g(\frac{p_j}{Q_k} t)) e^{-t} \, dt \\
&\geq C_1 \int_0^\infty \prod_{j=1}^{k-K} (1 - g(C_2 p^{-j} t)) e^{-t} \, dt, \text{ for } k \geq K,
\end{aligned}
$$

where C_1, C_2 are positive constants. Next,

$$
\prod_{j=1}^{k-K} (1 - g(C_2 p^{-j} t)) \geq \prod_{j=1}^{\infty} (1 - g(C_2 p^{-j} t)).
$$

Show that the latter is strictly positive:

$$
\sum_{j=1}^{\infty} g(C_2 t p^{-j}) = \sum_{j=1}^{\infty} e^{-C_2 t p^{-j}} (1 + C_2 t p^{-j}) \leq C_3 \sum_{j=1}^{\infty} e^{-C_4 t p^{-j}} < \infty
$$

for all t, thus $\prod_{j=1}^{k-K} (1 - g(C_2 p^{-k} t))$ is bounded away from zero uniformly in $k, k \geq K$, by $h(t) = \prod_{j=1}^{\infty} (1 - g(C_2 p^{-j} t)) > 0$, and $\mathbf{P}(B_k \mid A_k) \geq C_1 \int_0^\infty h(t) e^{-t} \, dt > 0$, so the series $\sum_{k=1}^{\infty} \mathbf{P}(B_k)$ diverges and the DB model is recurrent under the assumptions of (i).

(ii) Now, assume that $p_k \sim C e^{-\alpha k^\gamma}$. Consider the total time ν spent in the ground state, when none of the bits is active. We are going to prove for this particular choice of p_k that the expected time spent in the ground state

$$
\mathbf{E}\, \nu = \int_0^\infty \prod_{k=1}^{\infty} (1 - p_k t e^{-p_k t}) \, dt
$$

is finite. The product under the integral is bounded by

$$
\prod_{k=1}^{\infty} (1 - p_k t e^{-p_k t}) \leq \exp \Big\{ \operatorname{card}\{k : l_{1,\varepsilon} \leq p_k t \leq l_{2,\varepsilon}\} \log(1 - 1/e + \varepsilon) \Big\}.
$$

Here $l_{1,\varepsilon}, l_{2,\varepsilon}$ are the left and the right boundaries of the interval, where the function $x e^{-x}$ is greater or equal than $1/e - \varepsilon$. Taking into account the particular choice of p_k, we write:

$$
\begin{aligned}
\operatorname{card}\{k : l_{1,\varepsilon} \leq p_k t \leq l_{2,\varepsilon}\} &\sim \left(\frac{1}{\alpha} \log \frac{tC}{l_{1,\varepsilon}} \right)^{1/\gamma} - \left(\frac{1}{\alpha} \log \frac{tC}{l_{2,\varepsilon}} \right)^{1/\gamma} \\
&\sim \frac{\log l_{2,\varepsilon} - \log l_{1,\varepsilon}}{\gamma \alpha^{\frac{1}{\gamma} - 1}} (\log(tC))^{\frac{1}{\gamma} - 1}, \qquad (14)
\end{aligned}
$$

hence the infinite product in question is integrable for $\gamma < 1/2$. $\square$

Remark 3.1. *The sufficient condition in Theorem 2.4, (i) is slightly stronger than a condition similar to the one in Theorem 2.1, (i):*

$$\limsup_{k\to\infty} \beta^k p_k < \infty \text{ for some constant } \beta > 1. \tag{15}$$

It is not hard to see that the assumption of Theorem 2.4, (i) implies (15) *for $\beta = 1/(p+\varepsilon)$ for any $\varepsilon \in (0, 1-p)$. The converse implication is not true in general, for a counterexample we can put*

$$\kappa(k) = \min\{j^2 : j \in \mathbb{N} \text{ and } j^2 > k\},$$

and then

$$p_k = C2^{-\kappa(k)}, \quad k = 1, 2, \ldots$$

for a suitable constant C. Then (15) *holds with $\beta = 2$. The assumption of Theorem 2.4, (i) fails to hold: for the subsequence $k_i = i^2, i = 1, 2, \ldots$ we have*

$$\begin{aligned}\frac{Q_{k_i}}{Q_{k_i-1}} &= 1 - \frac{p_{k_i}}{Q_{k_i-1}} = 1 - \frac{p_{i^2}}{\sum_{j=k_i}^{\infty} p_j} \geq 1 - \frac{p_{i^2}}{\sum_{j=i^2}^{(i+1)^2-1} p_j} \\ &= 1 - \frac{C2^{-(i+1)}}{((i+1)^2 - i^2)C2^{-(i+1)}} \geq 1 - \frac{1}{2i+1} \to 1, \quad i \to \infty.\end{aligned}$$

However, the converse implication will hold if in addition to (15) *we require, for instance, the sequence $\{Q_k/Q_{k+1}\}$ to be monotone.*

3.3 The Central Limit Theorem

For the proof of the CLT for the number of active bits in BF and DB models we use the following general CLT for the triangular array, see, e.g., [4, Ch.8, Theorem 5]:

Theorem 3.3. *Let $\{\xi_{k,n}\}$, $1 \leq k \leq r_n$, $1 \leq n \leq \infty$ be a triangular array of random variables such that $\mathbf{E}\,\xi_{k,n} = 0$ and that the random variables $(\xi_{k,n})_{1\leq k\leq r_n}$ are mutually independent inside of every row $n = 1, 2, \ldots$. Assume that:*

(i) $\sum_{k=1}^{r_n} \mathbf{E}\,\xi_{k,n}^2 = 1,$

(ii) $\sum_{k=1}^{r_n} \mathbf{E}\left[\xi_{k,n}^2;\, |\xi_{k,n}| > M\right] \to 0,\ n \to \infty,$ *for every $M > 0$.*

Then

$$\sum_{k=1}^{r_n} \xi_{k,n} \overset{D}{\Longrightarrow} \mathcal{N}(0,1), \text{ as } n \to \infty$$

Proof of Theorem 2.5. It is easy to see that the expected number of active bits $\mathbf{E}\,\eta_t$ in BF model tends to infinity. We can write $\mathbf{E}\,\eta_t$ explicitly as

$$\mathbf{E}\,\eta_t = \sum_{k=1}^{\infty} \mathbf{P}\{\zeta_t^k = 1\} = \sum_{k=1}^{\infty} \frac{1}{2}(1 - e^{-2p_k t}).$$

Every term in the latter sum monotonously approaches $1/2$ as $t \to \infty$, thus the whole sum tends to infinity.

Next, for the DB model, given the assumption (4), if we fix a small $\varepsilon > 0$ and take $l_{1,\varepsilon}, l_{2,\varepsilon}$ to be as in (14) the left and the right borders of the interval where the function xe^{-x} is greater than $1/e - \varepsilon$, then, by the same reasoning as in (14), we obtain:

$$\begin{aligned}\mathbf{E}\,\eta_t &= \sum_{k=1}^{\infty} \mathbf{P}\{\zeta_t^k = 1\} = \sum_{k=1}^{\infty} p_k t e^{-p_k t} \\ &\geq (e^{-1} - \varepsilon)\,\mathrm{card}\{k : \lambda_{1,\varepsilon} \leq p_k t \leq \lambda_{2,\varepsilon}\} \geq C_1(\log(tC))^{\frac{1}{\gamma}-1} \to \infty\end{aligned}$$

for a constant C_1 depending on ε, γ and α.

The rest of the proof works for both BF and DB models. It is sufficient to prove the CLT for the embedded discrete time process $\{\eta_{T_n}\}_{n\geq 1}$ for an arbitrary non-random time sequence $\{T_n\}_{n\geq 1}$ going to infinity. Let us fix such a sequence and denote $\zeta_n := \zeta_{T_n}$ and $\eta_n := \eta_{T_n}$, for short. Introduce random variables

$$\begin{aligned} Z_{n,k} &= \mathbb{I}\{\zeta_n^k = 1\}, \\ \xi_{n,k} &= \begin{cases} \frac{Z_{n,k} - \mathbf{E}\,Z_{n,k}}{\sqrt{\mathbf{var}\,\eta_n}}, & k < r_n, \\ \frac{\sum\limits_{k \geq r_n} (Z_{n,k} - \mathbf{E}\,Z_{n,k})}{\sqrt{\mathbf{var}\,\eta_n}}, & k = r_n. \end{cases}\end{aligned}$$

We leave ourselves a freedom to choose a suitable sequence $\{r_n\}$ later. Check the conditions of Theorem 3.3. The random variables $\{\xi_{n,k}\}_{k=1}^{r_n}$ are mutually independent for every n. Condition (i) holds trivially. As for (ii), one has:

$$\begin{aligned}\sum_{1\leq k\leq r_n} \mathbf{E}\left[\xi_{n,k}^2;\, |\xi_{n,k}| > M\right] &= \overbrace{\sum_{1\leq k\leq r_n - 1} \mathbf{E}\left[\xi_{n,k}^2;\, |\xi_{n,k}| > M\right]}^{S_1} \\ &\quad + \underbrace{\mathbf{E}\left[\xi_{n,r_n}^2;\, |\xi_{r_n,n}| > M\right]}_{S_2}. \end{aligned} \tag{16}$$

By the assumptions $\mathbf{E}\,\eta(t) \to \infty$ as $t \to \infty$. Moreover,

$$C_2 \mathbf{E}\,\eta(t) \leq \mathbf{var}\,\eta(t) = \sum_{k\geq 1} f(p_k t)(1 - f(p_k t)) \leq \mathbf{E}\,\eta(t),$$

where $f(x) = \frac{1}{2}(1 - e^{-x})$ in BF model, $f(x) = xe^{-x}$ in DB model, and $C_2 = (1 - \sup_{x \in \mathbb{R}^+} f(x))$, with the respective f, so that $0 < C_2 < 1$ in both cases. Hence, by construction of $\xi_{n,k}$, the sum S_1 in (16) tends to 0 as n goes to infinity, because almost surely $\xi_{n,k} \leq 1/\mathbf{var}\, \eta_n \to 0$ and every term in S_1 is eventually zero. Lastly,

$$\mathbf{E}\, \xi_{r_n,n}^2 = \frac{1}{\mathbf{var}\, \eta_n} \sum_{k \geq r_n} f(p_k T_n)(1 - f(p_k T_n))$$

and so we can choose such r_n that the latter sum is no larger than, for instance, $\sqrt{\mathbf{var}\, \eta_n}$, thus satisfying Condition (ii) of Theorem 3.3 and finishing the proof. □

Acknowledgements

The authors thank Sergey Foss for the discussions from which the Bit Flipping models we consider here started, as well as for the follow-up talks and insights on relation of Bit Flipping to other fields. The authors are grateful to Robin Pemantle for an idea of a continuous-time implementation of the process, which proved to be an irreplaceable tool in the analysis. We are grateful to two anonymous referees for thorough reading and their valuable comments which allowed us to significantly improve the presentation of the material.

References

[1] S. Aspandiiarov, R. Iasnogorodski, and M. Menshikov. Passage-time moments for nonnegative stochastic processes and an application to reflected random walks in a quadrant. *Ann. Probab.*, 24:932–960, 1996.

[2] K. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching process and related limit theorem. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[3] J. Balogh and R. Pemantle. The Klee-Minty random edge chain moves with linear speed. *Random Structures and Algorithms*, 30(4):371–390, 2007.

[4] A. Borovkov. *Probability Theory.* Amsterdam : Gordon and Breach, 1998.

[5] A. Deza, E. Nematollahi, and T. Terlaky. How good are interior point methods? Klee-Minty cubes tighten iteration-complexity bounds. *Math. Program., Ser. A*, 113(1):1–14, 2008.

[6] R. Lyons, R. Pemantle, and Y. Peres. Random walks on the lamplighter group. *Ann. Probab*, 24:1993–2006, 1996.

[7] J. Steif. A survey of dynamical percolation. In *Fractal geometry and stochastics IV*, pages 145–174. Springer, 2009.

[8] W. Woess. *Random walks on infinite graphs and groups*, volume 138. Cambridge university press, 2000.