

Mixing of the noisy voter model

Harishchandra Ramadas*

Abstract

We prove that the noisy voter model mixes extremely fast – in time of $O(\log n)$ on any graph with n vertices – for arbitrarily small values of the ‘noise parameter’. We then explain why, as a result, this is an example of a spin system that is *always* in the ‘high-temperature regime’.

Keywords: noisy voter model; spin system; mixing time.

AMS MSC 2010: 82C20; 82B20.

Submitted to ECP on August 16, 2013, final version accepted on March 6, 2014.

1 Introduction

The *noisy voter model* is a spin system that can serve as a crude stochastic model for the spread of opinions in a human population, or the competition for territory between two species. The model was introduced by Granovsky and Madras [4] as a simple ergodic variant of the extensively-studied *voter model* [7]. Our main result in this paper (Theorem 3.1) is an upper bound on the mixing time of this model: we show that on an arbitrary graph with n vertices, the system mixes extremely fast – in time of $O(\log n)$ – for *all* values of the ‘noise parameter’, δ .

It is well-known that fast *temporal* mixing is equivalent to a *spatially* ‘well-mixed’ stationary measure [3]. Corollary 4.2 states that the stationary measure of the noisy voter model on any subset of the integer lattice satisfies a decay-of-correlations condition called *strong spatial mixing*. Conventional spin models such as the Ising model, gas hard-core model and proper coloring model exhibit strong spatial mixing only when they are sufficiently ‘disordered’, e.g. the Ising model on boxes in \mathbb{Z}^2 above the critical temperature. The noisy voter model is thus extremely special, since it can be thought of as *always* being in the high-temperature regime.

2 Preliminaries

2.1 The Noisy Voter Model

Starting with a finite simple undirected graph G with vertex set $V(G)$, the state space we consider is the set of all 0-1 configurations on $V(G)$, denoted by $\{0, 1\}^{V(G)}$. In other words, at each site $v \in V(G)$, we place a ‘spin’, which may have value 0 or 1. For $\eta \in \{0, 1\}^{V(G)}$ and $v \in V(G)$, we write $\eta(v)$ for the value of the spin at the site v . For $x \in V(G)$, we denote by η^x the configuration obtained from η by flipping the spin at x , leaving all other spins unchanged:

*University of Washington, Seattle, USA. E-mail: ramadas@math.washington.edu

$$\eta^x(v) := \begin{cases} 1 - \eta(v) & v = x, \\ \eta(v) & v \neq x. \end{cases}$$

Further, we denote by $d(x)$ the degree of the vertex x and write $y \sim x$ to indicate that the vertices x and y are neighbours.

Definition 2.1. Given a finite graph G and a function

$$c : V(G) \times \{0, 1\}^{V(G)} \rightarrow [0, \infty),$$

a spin system is a continuous-time Markov chain on the state space $\{0, 1\}^{V(G)}$ with transition-rate matrix $q(\cdot, \cdot)$, where

$$q(\eta, \eta') = \begin{cases} c(x, \eta) & \text{if } \eta' = \eta^x \text{ for some } x \in V(G), \\ 0 & \text{otherwise,} \end{cases}$$

for all configurations $\eta, \eta' \in \{0, 1\}^{V(G)}$ such that $\eta \neq \eta'$.

The voter model, introduced in the 70s, is an extremely well-understood spin system [7, Chapter V]. It is defined by the transition rates

$$c(x, \eta) = \frac{1}{d(x)} \#\{y : y \sim x, \eta(y) \neq \eta(x)\},$$

i.e., the rate of flipping the spin at x while in state η is proportional to the number of neighbours of x with spin differing from that at x .

We imagine a ‘voter’ located at each vertex x and that the spin at x corresponds to its ‘opinion’, either 0 or 1. Another interpretation of the voter model is a situation where two species, denoted by 0 and 1, compete for territory.

Intuitively, we imagine independent rate-1 Poisson clocks located at each vertex $v \in V(G)$ (i.e., clocks that ring after independent $\text{Exp}(1)$ waiting times). If the clock at vertex x rings, the voter at x picks one of its neighbours uniformly at random, and modifies its own opinion to match that of the chosen neighbour.

We observe, however, that the voter model is *not* ergodic, since the all-0 and all-1 states (denoted $\underline{0}$ and $\underline{1}$ respectively) are absorbing. There are hence multiple invariant measures: $\delta_{\underline{0}}$, $\delta_{\underline{1}}$ and all their convex combinations.

The model we study in this paper is a simple ergodic variant of the voter model – the **noisy voter model**, introduced by Granovsky and Madras in 1995 [4].

Definition 2.2. Let $\delta > 0$. The noisy voter model is a spin system on a graph G with transition rates

$$c(x, \eta) = \frac{1}{2\delta + 1} \left[\frac{1}{d(x)} \#\{y : y \sim x, \eta(y) \neq \eta(x)\} + \delta \right], \quad (2.1)$$

where $d(x)$ is the degree of the vertex x .

The factor $\frac{1}{2\delta+1}$ in (2.1) ‘normalises’ the system so that we have the following intuitive description of the dynamics: assign independent rate-1 Poisson clocks to each vertex $v \in V(G)$. Every time a clock rings at a vertex v , erase $\eta(v)$ and set

- $\eta(v) = 0$ with probability $\frac{1}{2\delta+1} \left[\frac{1}{d(v)} \#\{y : y \sim v, \eta(y) = 0\} + \delta \right]$,
- $\eta(v) = 1$ with probability $\frac{1}{2\delta+1} \left[\frac{1}{d(v)} \#\{y : y \sim v, \eta(y) = 1\} + \delta \right]$.

The parameter δ is called the ‘noise’ for obvious reasons, and we assume that it is strictly positive. (The case $\delta = 0$ is just the regular voter model, which, as explained above, is not ergodic.)

Remark 2.3. We mention briefly that this is not the most general form of the noisy voter model – we could alternatively have two positive parameters δ and β so that the noise in the model causes the system to preferentially tend towards spin 0 or spin 1. In this case, the flip rates can be written as follows:

$$c(x, \eta) = \frac{1}{\delta + \beta + 1} \left[\frac{1}{d(x)} \#\{y : y \sim x, \eta(y) \neq \eta(x)\} + \delta \mathbf{1}_{\eta(x)=0} + \beta \mathbf{1}_{\eta(x)=1} \right].$$

However, for the sake of simplicity of the proof, we choose to restrict ourselves to a single noise parameter, δ ; it can be easily verified that the proof goes through even in the more general case.

2.2 Mixing Time

We first formally define the intuitive notion of ‘distance to equilibrium’ for an ergodic finite-state Markov chain, or more generally, the distance between two probability measures on a finite set.

Definition 2.4. For any two probability measures μ_1 and μ_2 on a finite set S , the total variation distance between μ_1 and μ_2 is defined to be

$$\|\mu_1 - \mu_2\| := \max_{A \subseteq S} |\mu_1(A) - \mu_2(A)|.$$

We define next the notion of the mixing time of a Markov Chain (as a function of ϵ), the minimum time it takes to get to within a total variation distance of ϵ to stationarity:

Definition 2.5. For $0 < \epsilon < 1$, the (total variation) mixing time of an ergodic continuous-time Markov chain $(X_t)_{t \geq 0}$ on a finite state space S with stationary distribution μ is defined to be

$$t_{\text{mix}}(\epsilon) := \inf\{t \geq 0 : \max_{x \in S} \|\mathbb{P}_x(X_t \in \cdot) - \mu\| \leq \epsilon\}.$$

For two random variables X and Y with values in the same state space, we write $\|X - Y\|$ for the total variation distance between their distributions.

Definition 2.6. We say that a (continuous-time) spin system has optimal temporal mixing if there exist constants $C, c > 0$ such that the dynamics on any finite graph G has the following property: for any two copies $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ of the chain with possibly different initial configurations, we have

$$\|X_t - Y_t\| \leq Cne^{-ct}, \tag{2.2}$$

where, as mentioned above, the left hand side refers to the total variation distance between the distributions of the $\{0, 1\}^{V(G)}$ -valued random variables X_t and Y_t , and $n = |V(G)|$, the number of vertices in G .

Remark 2.7. We observe that optimal temporal mixing is equivalent to a mixing time of $O\left(\log\left(\frac{n}{\epsilon}\right)\right)$:

- Assume first that optimal temporal mixing holds, i.e., for two instances $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ of the dynamics, $\|X_t - Y_t\| \leq Cne^{-ct}$. We further assume that $Y_0 \stackrel{d}{=} \mu$, where μ is the equilibrium measure of the dynamics. Hence $Y_t \stackrel{d}{=} \mu$ for all $t \geq 0$. We thus obtain $\|X_t - \mu\| \leq Cne^{-ct}$. Now, suppose $t = \frac{1}{c} \log\left(\frac{nC}{\epsilon}\right) = O\left(\log\left(\frac{n}{\epsilon}\right)\right)$. Then, clearly $\|X_t - \mu\| \leq \epsilon$.

- Conversely, assume that $t_{\text{mix}}(\epsilon) \leq C \log\left(\frac{n}{\epsilon}\right)$, and let $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ be two instances of the dynamics. Then, for all $t > 0$, $t = C \log\left(\frac{2n}{\epsilon}\right)$ implies $\|X_t - \mu\| \leq \frac{\epsilon}{2}$. But $t = C \log\left(\frac{2n}{\epsilon}\right) \Leftrightarrow \epsilon = 2ne^{-t/C}$. Hence, for all $t > 0$,

$$\|X_t - Y_t\| \leq \|X_t - \mu\| + \|Y_t - \mu\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon = 2ne^{-t/C},$$

i.e., optimal temporal mixing holds.

Throughout the rest of this article, we keep the parameter ϵ fixed.

3 Main Result: Optimal Temporal Mixing

Theorem 3.1. For any $\delta > 0$, the noisy voter model on any finite graph G with n vertices has mixing time of $O(\log n)$.

Proof. A standard method for obtaining an upper bound on the total variation distance between the distributions of two random variables X and Y is to construct a suitable coupling (X, Y) , and to use the well-known inequality,

$$\|X - Y\| \leq \mathbb{P}(X \neq Y), \tag{3.1}$$

which holds for any coupling (X, Y) of X and Y (see, e.g. [9, Proposition 4.7]).

Further, if $(X_t, Y_t)_{t \geq 0}$ is a coupled spin system defined on a finite graph G , then for a fixed time t , a simple union bound yields

$$\begin{aligned} \mathbb{P}(X_t \neq Y_t) &= \mathbb{P}\left(\bigcup_{v \in V(G)} \{X_t(v) \neq Y_t(v)\}\right) \\ &\leq \sum_{v \in V(G)} \mathbb{P}(X_t(v) \neq Y_t(v)) \\ &\leq n \left(\max_{v \in V(G)} \mathbb{P}(X_t(v) \neq Y_t(v))\right). \end{aligned} \tag{3.2}$$

Hence, from (2.2), (3.1) and (3.2), it is enough to show that there exists a coupling $(X_t, Y_t)_{t \geq 0}$ such that for arbitrary starting configurations X_0 and Y_0 , the following holds for all $t \geq 0$:

$$\max_{v \in V(G)} \mathbb{P}(X_t(v) \neq Y_t(v)) \leq Ce^{-ct},$$

where C and c are fixed constants independent of G .

The Coupling: Since the noisy voter model is a monotone system, it suffices to consider a coupling (X_t, Y_t) of the system, with $X_0 = \underline{0}$ and $Y_0 = \underline{1}$. Place independent rate-1 Poisson clocks at each vertex $v \in V(G)$. When the clock at v rings:

- With probability $\frac{\delta}{2\delta+1}$, set both $X_t(v), Y_t(v) = 0$.
- With probability $\frac{\delta}{2\delta+1}$, set both $X_t(v), Y_t(v) = 1$.
- With probability $\frac{1}{2\delta+1}$, pick a neighbour u of v uniformly at random. Set $X_t(v) = X_t(u)$ and $Y_t(v) = Y_t(u)$.

It is easy to see that both X_t and Y_t follow the dynamics of the noisy voter model, defined in (2.1).

Let $Z_t(v) := \mathbf{1}_{\{X_t(v) \neq Y_t(v)\}}$ be the *disagreement process*, i.e., the indicator of the event that the processes starting from the $\underline{0}$ and $\underline{1}$ configurations (X_t and Y_t respectively) disagree at vertex v at some time $t \geq 0$. Obviously, $Z_0(v) = 1$ for all $v \in V(G)$.

We hence wish to show that

$$\max_{v \in V(G)} \mathbb{P}(Z_t(v) = 1) \leq Ce^{-ct}.$$

With this coupling, every time the Poisson clock at vertex v rings, with probability at least $\frac{2\delta}{2\delta+1}$, both X_t and Y_t are updated at v with the same value (0 or 1). Hence, independent of its current value, $Z_t(v)$ is set to 0. With probability $\frac{1}{2\delta+1}$, the vertex v chooses a neighbour u uniformly at random. Independent of the current value of $Z_t(v)$, it is clear that if $Z_t(u) = 1$, then $Z_t(v)$ becomes 1 whereas if $Z_t(u) = 0$, then $Z_t(v)$ becomes 0.

Since $V(G)$ is a finite set, $(Z_t)_{t \geq 0}$ is just a continuous-time Markov chain on the (finite) state space $\{0, 1\}^{V(G)}$ with initial state $\underline{1}$. We denote the corresponding transition function $p_t(\cdot, \cdot)$ and transition-rate matrix $q(\cdot, \cdot)$. As before, for $\eta \in \{0, 1\}^{V(G)}$, denote by η^v the configuration with the value at v flipped, and everywhere else the same as η . We can hence write a formula for the transition rates similar to the noisy voter model (2.1):

$$q(\eta, \eta^v) = \frac{2\delta}{2\delta+1} \mathbf{1}_{\{\eta(v)=1\}} + \frac{1}{2\delta+1} \cdot \frac{1}{d(v)} \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}}.$$

Interchanging η and η^v , we get

$$q(\eta^v, \eta) = \frac{2\delta}{2\delta+1} \mathbf{1}_{\{\eta(v)=0\}} + \frac{1}{2\delta+1} \cdot \frac{1}{d(v)} \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) = \eta(v)\}}.$$

Since $q(\eta, \eta') = 0$ if η and η' differ at more than one site and since for any Markov chain $\sum_y q(x, y) = 0$, we have

$$\begin{aligned} q(\eta, \eta) &= - \sum_{v \in V(G)} q(\eta, \eta^v) \\ &= - \sum_{v \in V(G)} \left(\frac{2\delta}{2\delta+1} \mathbf{1}_{\eta(v)=1} + \frac{1}{2\delta+1} \cdot \frac{1}{d(v)} \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}} \right). \end{aligned}$$

We can now write the forward Kolmogorov equation (see, e.g. [8, Equation 2.17]) for the Markov chain (Z_t) , started at $\underline{1}$:

$$\begin{aligned} \frac{d}{dt} p_t(\underline{1}, \eta) &= \sum_{v \in V(G)} [p_t(\underline{1}, \eta^v) \cdot q(\eta^v, \eta)] + p_t(\underline{1}, \eta) \cdot q(\eta, \eta) \\ &= \frac{2\delta}{2\delta+1} \sum_{v \in V(G)} p_t(\underline{1}, \eta^v) \mathbf{1}_{\{\eta(v)=0\}} + \frac{1}{2\delta+1} \sum_{v \in V(G)} \left(\frac{1}{d(v)} \cdot p_t(\underline{1}, \eta^v) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) = \eta(v)\}} \right) \\ &\quad - \frac{2\delta}{2\delta+1} \sum_{v \in V(G)} p_t(\underline{1}, \eta) \mathbf{1}_{\{\eta(v)=1\}} - \frac{1}{2\delta+1} \sum_{v \in V(G)} \left(\frac{1}{d(v)} \cdot p_t(\underline{1}, \eta) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}} \right). \end{aligned} \tag{3.3}$$

Now, for $x \in V(G)$,

$$\frac{d}{dt} \mathbb{P}(Z_t(x) = 1) = \frac{d}{dt} \sum_{\eta \in \{0,1\}^{V(G)}} p_t(\underline{1}, \eta) \cdot \mathbf{1}_{\{\eta(x)=1\}}. \tag{3.4}$$

Substituting (3.3) in (3.4) above, with the convention that sums indexed with v and η run over all vertices and all configurations respectively,

$$\begin{aligned} \frac{d}{dt} \mathbb{P}(Z_t(x) = 1) &= \frac{2\delta}{2\delta+1} \sum_{v, \eta} (p_t(\underline{1}, \eta^v) \mathbf{1}_{\{\eta(v)=0, \eta(x)=1\}} - p_t(\underline{1}, \eta) \mathbf{1}_{\{\eta(v)=1, \eta(x)=1\}}) \\ &\quad + \frac{1}{2\delta+1} \sum_{v, \eta} \frac{1}{d(v)} \cdot \mathbf{1}_{\{\eta(x)=1\}} \cdot \left(p_t(\underline{1}, \eta^v) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) = \eta(v)\}} - p_t(\underline{1}, \eta) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}} \right). \end{aligned} \tag{3.5}$$

For each of the sums over v in (3.5) above, we consider the cases $v = x$ and $v \neq x$ separately to get

$$\begin{aligned} & \frac{d}{dt} \mathbb{P}(Z_t(x) = 1) \\ &= \frac{2\delta}{2\delta + 1} \sum_{\eta} (p_t(\underline{1}, \eta^x) \cdot 0 - p_t(\underline{1}, \eta) \cdot \mathbf{1}_{\{\eta(x)=1\}}) \\ &+ \frac{2\delta}{2\delta + 1} \sum_{v:v \neq x} \sum_{\eta} (p_t(\underline{1}, \eta^v) \cdot \mathbf{1}_{\{\eta^v(v)=1, \eta^v(x)=1\}} - p_t(\underline{1}, \eta) \cdot \mathbf{1}_{\{\eta(v)=1, \eta(x)=1\}}) \\ &+ \frac{1}{2\delta + 1} \left\{ \sum_{\eta} \frac{1}{d(x)} \cdot \mathbf{1}_{\{\eta(x)=1\}} \cdot \left(p_t(\underline{1}, \eta^x) \sum_{u:u \sim x} \mathbf{1}_{\{\eta(u)=1\}} - p_t(\underline{1}, \eta) \sum_{u:u \sim x} \mathbf{1}_{\{\eta(u)=0\}} \right) \right. \\ &\left. + \sum_{v:v \neq x} \frac{1}{d(v)} \sum_{\eta} \mathbf{1}_{\{\eta(x)=1\}} \cdot \left(p_t(\underline{1}, \eta^v) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u)=\eta(v)\}} - p_t(\underline{1}, \eta) \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}} \right) \right\}. \end{aligned}$$

In the second term on the right hand side, the sum over η is clearly zero. Since in the second term in the braces $\{\dots\}$, $v \neq x$, it can be rewritten as

$$\begin{aligned} & \sum_{v:v \neq x} \frac{1}{d(v)} \sum_{\eta} \left(p_t(\underline{1}, \eta^v) \cdot \mathbf{1}_{\{\eta^v(x)=1\}} \sum_{u:u \sim v} \mathbf{1}_{\{\eta^v(u) \neq \eta^v(v)\}} \right. \\ & \left. - p_t(\underline{1}, \eta) \cdot \mathbf{1}_{\{\eta(x)=1\}} \sum_{u:u \sim v} \mathbf{1}_{\{\eta(u) \neq \eta(v)\}} \right), \end{aligned}$$

and as before, the sum over η equals zero. We are hence left with

$$\begin{aligned} & \frac{d}{dt} \mathbb{P}(Z_t(x) = 1) \\ &= \frac{2\delta}{2\delta + 1} \sum_{\eta} (-p_t(\underline{1}, \eta) \cdot \mathbf{1}_{\{\eta(x)=1\}}) \\ &+ \frac{1}{2\delta + 1} \sum_{\eta} \frac{1}{d(x)} \cdot \mathbf{1}_{\{\eta(x)=1\}} \cdot \left(p_t(\underline{1}, \eta^x) \sum_{u:u \sim x} \mathbf{1}_{\{\eta(u)=1\}} - p_t(\underline{1}, \eta) \sum_{u:u \sim x} \mathbf{1}_{\{\eta(u)=0\}} \right). \end{aligned}$$

We finally write the transition function elements and indicator functions in terms of probabilities to get the following ‘master equation’, valid for all $x \in V(G)$:

$$\begin{aligned} \frac{d}{dt} \mathbb{P}(Z_t(x) = 1) &= \frac{-2\delta}{2\delta + 1} \mathbb{P}(Z_t(x) = 1) \\ &- \frac{1}{2\delta + 1} \cdot \frac{1}{d(x)} \left(\sum_{u:u \sim x} \mathbb{P}(Z_t(x) = 1, Z_t(u) = 0) \right) \\ &+ \frac{1}{2\delta + 1} \cdot \frac{1}{d(x)} \left(\sum_{u:u \sim x} \mathbb{P}(Z_t(x) = 0, Z_t(u) = 1) \right). \end{aligned}$$

For notational simplicity, we now define

$$M(t) := \max_{1 \leq k \leq n} \mathbb{P}(Z_t(x_k) = 1),$$

where x_1, \dots, x_n are the n vertices of G . Suppose, at some instant $s \geq 0$,

$$M(s) = \mathbb{P}(Z_s(x_k) = 1),$$

for some k with $1 \leq k \leq n$. Then $\mathbb{P}(Z_s(x_k) = 1) \geq \mathbb{P}(Z_s(x_l) = 1)$ for all $l \neq k$, $1 \leq l \leq n$. Hence,

$$\begin{aligned} \mathbb{P}(Z_s(x_k) = 1, Z_s(x_l) = 0) &= \mathbb{P}(Z_s(x_k) = 1) - \mathbb{P}(Z_s(x_k) = 1, Z_s(x_l) = 1) \\ &\geq \mathbb{P}(Z_s(x_l) = 1) - \mathbb{P}(Z_s(x_k) = 1, Z_s(x_l) = 1) \\ &= \mathbb{P}(Z_s(x_k) = 0, Z_s(x_l) = 1). \end{aligned}$$

Setting $t = s$ and $x = x_k$ in (3.1), we see that the magnitude of the second term on the right hand side is greater than that of the third term, since for all neighbours u of x_k ,

$$\mathbb{P}(Z_s(x_k) = 1, Z_s(u) = 0) \geq \mathbb{P}(Z_s(x_k) = 0, Z_s(u) = 1).$$

As a result,

$$\begin{aligned} \frac{d}{dt} \mathbb{P}(Z_t(x_k) = 1) \Big|_{t=s} &\leq -\frac{2\delta}{2\delta + 1} \mathbb{P}(Z_s(x_k) = 1) \\ &= -\frac{2\delta}{2\delta + 1} M(s). \end{aligned} \tag{3.6}$$

In general, at the instant $t = s$, there might be multiple k ($1 \leq k \leq n$), such that $M(s) = \mathbb{P}(Z_s(x_k) = 1)$. However, (3.6) holds for every such k . Furthermore, $\mathbb{P}(Z_t(x_l) = 1)$ is a differentiable (hence continuous) function of t for each $l = 1, \dots, n$. Consequently, for all $t \geq 0$, we have the inequality

$$D^+ M(t) := \limsup_{h \rightarrow 0^+} \frac{M(t+h) - M(t)}{h} \leq -\frac{2\delta}{2\delta + 1} M(t),$$

where $D^+ M(t)$ is the *upper right Dini derivative* of M at t . Again, since $\mathbb{P}(Z_t(x_l) = 1)$ is continuous for each l , $M(t)$ is continuous at every $t \geq 0$, and we can use the fundamental theorem of calculus to write

$$M(t) = \frac{d}{dt} \int_0^t M(s) ds = D^+ \left(\int_0^t M(s) ds \right).$$

Hence,

$$D^+ M(t) + \frac{2\delta}{2\delta + 1} D^+ \left(\int_0^t M(s) ds \right) \leq 0.$$

Using the elementary fact that $D^+(f + g) \leq D^+ f + D^+ g$ for any two functions f and g , we have

$$D^+ \left(M(t) + \frac{2\delta}{2\delta + 1} \int_0^t M(s) ds \right) \leq 0,$$

for all $t \geq 0$. The monotonicity theorem for Dini derivatives [10, Appendix I] says that if $D^+ f(t) \leq 0$ for all $t \geq 0$, f must be nonincreasing. As a result,

$$M(t) + \frac{2\delta}{2\delta + 1} \int_0^t M(s) ds \leq M(0) = 1,$$

and hence,

$$M(t) \leq 1 - \frac{2\delta}{2\delta + 1} \int_0^t M(s) ds$$

for all $t \geq 0$. Finally, applying the Gronwall-Bellman inequality [1], we get

$$M(t) \leq e^{-\frac{2\delta}{2\delta+1} t},$$

i.e.,

$$\max_{v \in V(G)} \mathbb{P}(Z_t(v) = 1) \leq e^{-\frac{2\delta}{2\delta+1} t}.$$

We thus conclude that optimal temporal mixing holds, or equivalently, by Remark 2.7, that the system mixes in time of $O(\log n)$. \square

Remark 3.2. *It is generally believed that for any spin system, $O(\log n)$ mixing is indeed ‘optimal’: by a coupon-collector argument, one imagines that in $o(\log n)$ time, not enough sites are updated for the system to equilibrate. However, it is shown in [5] that*

this argument can fail, but the authors then prove that any reversible spin system with ‘local’ interactions on a bounded-degree graph with n vertices necessarily has mixing time of $\Omega(\log n)$. Unfortunately, a simple application of Kolmogorov’s criterion [6, Chapter 1] shows that the noisy voter model is in general not reversible. An exception is the simple case of the 1-dimensional torus, $\mathbb{Z}/n\mathbb{Z}$, for which it is easy to see that the noisy voter model with parameter δ is in fact exactly the same process as the stochastic (ferromagnetic) Ising model with inverse temperature $\beta = \frac{1}{4} \log(1 + \delta^{-1})$.

4 Always Hot: Strong Spatial Mixing

As mentioned in the introduction, it is already well-known that fast temporal mixing is in some sense equivalent to a ‘well-mixed’ stationary measure. We formalise this below.

Although our main result (optimal temporal mixing) holds for any graph on n vertices, we focus on the special case of ‘boxes’ in the d -dimensional integer lattice \mathbb{Z}^d , for some $d \geq 1$. (We do this for simplicity; in general, one could consider any lattice with subexponential growth.) By a (finite) *box* in \mathbb{Z}^d , we mean a finite induced subgraph: a finite subset of the vertex set, along with all edges between these vertices inherited from the lattice. (In particular, a box need not be cuboidal.)

For $u, v \in \mathbb{Z}^d$, the *distance* between u and v is defined to be

$$\text{dist}(u, v) := \sum_{i=1}^d |u_i - v_i|.$$

The distance between a vertex u and a box G is

$$\text{dist}(u, G) := \min_{v \in V(G)} \text{dist}(u, v).$$

For a box $G \subseteq \mathbb{Z}^d$, we define its *boundary* to be the set of vertices not in G that are neighbours with at least one vertex in G . A configuration $\tau \in \{0, 1\}^{\partial G}$ is said to be a *boundary condition* for the spin system dynamics on the box G , if we run the dynamics on $G \cup \partial G$ conditioned on the spins at the boundary being frozen at their initial values.

Finally, for a finite box $H \subseteq G \subset \mathbb{Z}^d$ and a boundary condition $\tau \in \{0, 1\}^{\partial G}$, we define $\mu_G^\tau|_H$ to be the stationary distribution for the dynamics run on G with boundary condition τ , *projected onto* $\{0, 1\}^{V(H)}$.

Dyer, Sinclair, Vigoda and Weitz give in [3] an elegant combinatorial proof of the fact that if a spin system has optimal temporal mixing, then the stationary distribution satisfies a condition called strong spatial mixing, which we define below.

Definition 4.1. *A spin system has strong spatial mixing if there exist constants $C, c > 0$ such that for any finite box $G \subseteq \mathbb{Z}^d$ and a box $H \subseteq G$, any site $u \in \partial G$ and any pair of boundary conditions τ and τ^u that differ only at u ,*

$$\|\mu_G^\tau|_H - \mu_G^{\tau^u}|_H\| \leq C|V(H)| \exp(-c \cdot \text{dist}(u, H)) \tag{4.1}$$

holds.

Due to [3, Theorem 2.3], we hence immediately have the following corollary of Theorem 3.1:

Corollary 4.2. *The noisy voter model has strong spatial mixing for all values of the noise parameter δ .*

For other spin systems like the Ising model (with zero magnetic field), proper coloring model and gas hard-core model, strong spatial mixing can be shown to hold only in certain ranges of the parameter space: in the Ising model, the temperature must be high enough; in the proper coloring model, the number of colours should be large enough; and in the gas hard-core model, the fugacity must be low enough. We hence see that the noisy voter model is extremely special, since it can be thought of as *always* being in the high-temperature, highly disordered regime.

Remark 4.3. *We also mention that the noisy voter model, when defined on the entire lattice \mathbb{Z}^d , is ergodic for all $\delta > 0$, an easy consequence of Dobrushin's ' $M < \epsilon$ ' criterion [2]. This in particular means that the noisy voter model always has a unique infinite-volume invariant measure, again in contrast to other spin systems like the Ising model with zero magnetic field ($d \geq 2$): there exists a unique infinite-volume Gibbs measure only above the critical temperature.*

Acknowledgments. This paper contains work done while writing my Master's thesis at the Ludwig-Maximilians-Universität and Technische Universität München, supported by the Elitenetzwerk Bayern. I thank my supervisor, Nina Gantert (TU München) for suggesting this topic, for the many enlightening discussions and for having taught me so much.

References

- [1] Richard Bellman, *The stability of solutions of linear differential equations*. Duke Math. J. 10 (4), 643-647, 1943. MR-0009408
- [2] R. L. Dobrushin, *Markov Processes with a Large Number of Locally Interacting Components: Existence of a Limit Process and Its Ergodicity*, Probl. Peredachi Inf., 7:2 (1971), 70-87. (In Russian. See instead [8, Section 4.2] for an explanation in English of the $M < \epsilon$ criterion.) MR-0310999
- [3] Martin Dyer, Alistair Sinclair, Eric Vigoda and Dror Weitz, *Mixing in Time and Space for Lattice Spin Systems: A Combinatorial View*. Randomization and Approximation Techniques in Computer Science, 6th International Workshop, RANDOM 2002. MR-2047027
- [4] Boris L. Granovsky and Neal Madras, *The Noisy Voter Model*. Stochastic Processes and their Applications, Volume 55, Issue 1, 23-43, January 1995. MR-1312146
- [5] Thomas P. Hayes and Alistair Sinclair, *A general lower bound for mixing of single-site dynamics on graphs*. Annals of Applied Probability, Volume 17, Number 3 (2007), 931-952. MR-2326236
- [6] Frank Kelly, *Reversibility and Stochastic Networks*. Wiley, 1979. MR-0554920
- [7] Thomas M. Liggett, *Interacting Particle Systems*. Springer, 2005 Reprint (Classics in Mathematics). MR-2108619
- [8] Thomas M. Liggett, *Continuous Time Markov Processes: An Introduction*. AMS, 2010. MR-2574430
- [9] David A. Levin, Yuval Peres, Elizabeth L. Wilmer, *Markov Chains and Mixing Times*. AMS, 2008. MR-2466937
- [10] Rouche, Nicolas, P. Habets, and M. LaLoy, *Stability Theory by Liapunov's Direct Method*. Springer, 1977. MR-0450715

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS¹)
- Non profit, sponsored by IMS², BS³, PKP⁴
- Purely electronic and secure (LOCKSS⁵)

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

²IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

³BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁴PK: Public Knowledge Project <http://pkp.sfu.ca/>

⁵LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>