Equilibrium interfaces of biased voter models

Rongfeng Sun [1] Jan M. Swart [2] Jinjiong Yu [3]

April 12, 2018

Abstract

A one-dimensional interacting particle system is said to exhibit interface tightness if starting in an initial condition describing the interface between two constant configurations of different types, the process modulo translations is positive recurrent. In a biological setting, this describes two populations that do not mix, and it is believed to be a common phenomenon in one-dimensional particle systems. Interface tightness has been proved for voter models satisfying a finite second moment condition on the rates. We extend this to biased voter models. Furthermore, we show that the distribution of the equilibrium interface for the biased voter model converges to that of the voter model when the bias parameter tends to zero. A key ingredient is an identity for the expected number of boundaries in the equilibrium voter model interface, which is of independent interest.

MSC 2010. Primary: 82C22; Secondary: 82C24, 82C41, 60K35.
Keywords. Biased voter model, interface tightness, branching and coalescing random walks.

Contents

[1] Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, 119076 Singapore. Email: matsr@nus.edu.sg
[2] The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod vodárenskou věží 4, 18208 Prague 8, Czech Republic. Email: swart@utia.cas.cz
[3] NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China. Email: jinjiongyu@nyu.edu

1 Introduction

1.1 Interface tightness

One-dimensional *biased voter models*, also known as one-dimensional Williams-Bjerknes models [Sch77, WB72], are Markov processes $(X_t)_{t\geq 0}$ taking values in the space $\{0,1\}^{\mathbb{Z}}$ of infinite sequences $x = (x(i))_{i\in\mathbb{Z}}$ of zeros and ones. They have several interpretations, one of which is to model the dynamics of two biological populations. We call $X_t(i)$ the type of the individual at site $i \in \mathbb{Z}$ at time $t \geq 0$. Let $\varepsilon \in [0,1)$, and let $a : \mathbb{Z} \to [0,\infty)$ be a function such that $\sum_k a(k) < \infty$ and the continuous-time random walk that jumps from i to j with rate $a(j-i)$ is irreducible. The dynamics of $(X_t)_{t\geq 0}$ are such that for each $i, j \in \mathbb{Z}$, at the times t of a Poisson point process with rate $a(j-i)$, if the type at i just prior to t satisfies $X_{t-}(i) = 1$, then j adopts the type 1; if $X_{t-}(i) = 0$, then with probability $1-\varepsilon$, site j adopts the type 0, and with the remaining probability ε, $X_t(j)$ remains unchanged.

Somewhat more formally, we can define $(X_t)_{t\geq 0}$ by specifying its generator. For $x \in \{0,1\}^{\mathbb{Z}}$ and $i_1,\ldots,i_n \in \mathbb{Z}$, write $x(i_1,\ldots,i_n) := (x(i_1),\ldots,x(i_n)) \in \{0,1\}^n$. We also use the convention of writing elements of $\{0,1\}^n$ as words consisting of the letters 0 and 1, i.e., instead of $(1,0)$ we simply write 10 and similarly for longer sequences. With this notation, the generator of the biased voter model we have just described is given by

$$\begin{aligned} G^\varepsilon f(x) = &\sum_{i,j} a(j-i) 1_{\{x(i,j)=10\}} \big\{ f(x+e_j) - f(x) \big\} \\ &+ (1-\varepsilon) \sum_{i,j} a(j-i) 1_{\{x(i,j)=01\}} \big\{ f(x-e_j) - f(x) \big\}, \end{aligned} \tag{1.1}$$

where $e_i(j) := 1_{\{i=j\}}$. We call ε the *bias* and a the *underlying random walk kernel*. In particular, for $\varepsilon = 0$, we obtain a normal (unbiased) voter model, in which the types 0 and 1 play symmetric roles. By contrast, for $\varepsilon > 0$, the 1's replace 0's at a faster rate than the other way round, i.e., there is a bias in favor of the 1's. To indicate the bias parameter, $(X_t)_{t\geq 0}$ will be denoted as $(X_t^\varepsilon)_{t\geq 0}$ hereafter.

Let

$$\begin{aligned} S^{01}_{\rm int} &:= \big\{ x \in \{0,1\}^{\mathbb{Z}} : \lim_{i\to-\infty} x(i) = 0,\ \lim_{i\to\infty} x(i) = 1 \big\}, \\ S^{10}_{\rm int} &:= \big\{ x \in \{0,1\}^{\mathbb{Z}} : \lim_{i\to-\infty} x(i) = 1,\ \lim_{i\to\infty} x(i) = 0 \big\} \end{aligned} \tag{1.2}$$

denote the sets of states describing the interface between an infinite cluster of 1's and an infinite cluster of 0's. If $\sum_k a(k)|k| < \infty$ and $0 \leq \varepsilon < 1$, then it is not hard to see that $X_0^\varepsilon \in S^{01}_{\rm int}$ implies that $X_t^\varepsilon \in S^{01}_{\rm int}$ for all $t \geq 0$, a.s., and similarly for $S^{10}_{\rm int}$. (For unbiased voter models, this is proved in [BMSV06]. The proof in the biased case is the same.) We will be interested in studying the long-time behavior of the interface of $(X_t^\varepsilon)_{t\geq 0}$.

We call two configurations $x, y \in \{0,1\}^{\mathbb{Z}}$ *equivalent*, denoted by $x \sim y$, if one is a translation of the other, i.e., there exists some $k \in \mathbb{Z}$ such that $x(i) = y(i+k)$ $(i \in \mathbb{Z})$. We let $\overline{x}$ denote the equivalence class containing x and write

$$\overline{S}^{01}_{\rm int} := \{\overline{x} : x \in S^{01}_{\rm int}\} \quad \text{and} \quad \overline{S}^{10}_{\rm int} := \{\overline{x} : x \in S^{10}_{\rm int}\}. \tag{1.3}$$

Note that $S^{01}_{\rm int}$, $\overline{S}^{01}_{\rm int}$, $S^{10}_{\rm int}$ and $\overline{S}^{10}_{\rm int}$ are countable sets. Since our rates are translation invariant, the *process modulo translations* $(\overline{X}^\varepsilon_t)_{t\geq 0}$ is itself a Markov process. For non-nearest neighbor kernels a, this Markov process is irreducible; see Lemma 2.1 below. From now on, we restrict ourselves to the non-nearest neighbor case. We adopt the following definition from [CD95].

Definition 1.1 (Interface tightness) *We say that $(X_t^\varepsilon)_{t\geq 0}$ exhibits interface tightness on $S^{01}_{\rm int}$ (resp. $S^{10}_{\rm int}$) if $(\overline{X}^\varepsilon_t)_{t\geq 0}$ is positive recurrent on $\overline{S}^{01}_{\rm int}$ (resp. $\overline{S}^{10}_{\rm int}$).*

Note that because of the bias, if a is not symmetric, then there is no obvious symmetry telling us that interface tightness on $S^{01}_{\rm int}$ implies interface tightness on $S^{10}_{\rm int}$ or vice versa.

In the unbiased case $\varepsilon = 0$, Cox and Durrett [CD95] proved that interface tightness holds if $\sum_k a(k)|k|^3 < \infty$. This was relaxed to $\sum_k a(k)k^2 < \infty$ by Belhaouari, Mountford and Valle [BMV07], who moreover showed that the second moment condition is optimal. Our first main result is the following theorem that extends this to biased voter models, where the optimal condition in the biased case turns out to be even weaker.

Theorem 1.2 (Interface tightness for biased voter models) *Assume that $\varepsilon \in [0,1)$ and that the kernel a is non-nearest neighbor, irreducible and satisfies $\sum_k a(k)k^2 < \infty$. Then the biased voter model $(X^\varepsilon_t)_{t\geq 0}$ exhibits interface tightness on $S^{01}_{\rm int}$ and $S^{10}_{\rm int}$. If $\varepsilon > 0$ and the moment condition on the kernel is relaxed to $\sum_{k<0} a(k)k^2 < \infty$ and $\sum_{k>0} a(k)k < \infty$, then interface tightness on $S^{01}_{\rm int}$ still holds.*

To see heuristically why for interface tightness on $S^{01}_{\rm int}$, a finite first moment condition in the positive direction suffices, we observe that $a(k)$ with large positive k govern 0's that are created deep into the territory of the 1's. Such 0's do not survive long due to the bias. On the other hand, $a(k)$ with large negative k govern 1's that are created deep into the territory of the 0's. These 1's have a positive probability of surviving and then giving birth to 1's even further away. This explains heuristically why one needs to impose a stronger moment condition in the negative direction.

The proof of interface tightness for the voter model in [CD95, BMV07] relied heavily on the well-known duality of the voter model to coalescing random walks. A biased voter model also has a dual, which is a system of coalescing random walks that moreover branch with rate ε. Because of the branching, this dual process is much harder to control than in the unbiased case. In view of this, we were unable to apply the methods of [CD95, BMV07] but instead adapted a method of [SS08b], who provided a short proof of interface tightness for unbiased voter models using generator calculations and a Lyapunov like function. Our key observation is that this function admits a generalization to biased voter models, and the proof of interface tightness can then be adapted accordingly. However, as we will see in the proof of Theorem 1.2, the calculations are considerably more complicated in the biased case.

Interface tightness is believed to be a common phenomenon in one-dimensional interacting particle systems. Besides the voter model, it has also been proved for one-dimensional multi-type contact processes, see e.g., [AMPV10] and [Val10], where the proofs rely on duality, renormalization and percolation techniques.

1.2 Convergence of equilibrium interface

Theorem 1.2 implies that, modulo translations, the biased voter model is an irreducible countable-state Markov chain, and hence has a unique invariant law and is ergodic. In particular, if $\sum_k a(k)k^2 < \infty$, then for each $\varepsilon \geq 0$, there is a unique invariant law $\overline{\nu}_\varepsilon$ on $\overline{S}^{01}_{\rm int}$. It is a natural question whether $\overline{\nu}_\varepsilon$ converges to the unique invariant law $\overline{\nu}_0$ for the voter model. Our second main result answers this question affirmatively.

Theorem 1.3 (Continuity of the invariant law) *Assume that the kernel a is non-nearest neighbor, irreducible, and satisfies $\sum_k a(k)k^2 < \infty$. Then as $\varepsilon \downarrow 0$, the invariant law $\overline{\nu}_\varepsilon$ converges weakly to $\overline{\nu}_0$ with respect to the discrete topology on $\overline{S}^{01}_{\rm int}$.*

The proof of Theorem 1.3 turns out to be much more delicate than the result may suggest. The difficulty lies in the choice of the discrete topology on $\overline{S}^{01}_{\rm int}$. In particular, Theorem 1.3 implies that the length of the equilibrium interface under $\overline{\nu}_\varepsilon$ is tight as $\varepsilon \downarrow 0$, and if started at the *heaviside state* x_0 with

$$x_0(i) = 0 \quad \text{for } i < 0 \qquad \text{and} \qquad x_0(i) = 1 \quad \text{for } i \geq 0, \tag{1.4}$$

then the time it takes to return state x_0 is also tight. Such uniform control in ε as $\varepsilon \downarrow 0$ turns out to be difficult to obtain. We get around this difficulty by first proving the weak convergence of $\overline{\nu}_\varepsilon$ to $\overline{\nu}_0$ under the product topology, where we identify $\overline{S}^{01}_{\rm int}$ with a subset of the product space $\{0,1\}^{\mathbb{N}}$ by shifting the leftmost 1 of the interface to the origin. To strengthen the convergence to the discrete topology on $\overline{S}^{01}_{\rm int}$, we then establish uniform control (w.r.t. ε) on the expected number of boundaries in the equilibrium biased voter interface, as well as an exact formula for the expected number of such boundaries in the equilibrium voter model interface (see Proposition 3.7). This last result is also of independent interest.

1.3 Relation to the Brownian net

Theorem 1.3 is in fact motivated by studies of branching-coalescing random walks and their convergence to the Brownian net under weak branching. Let us now explain this connection.

Similar to the well-known duality between the voter model and coalescing random walks, the biased voter models are dual to systems of branching-coalescing random walks, with bias ε being the branching rate of the random walks. While in [CD95, BMV07], coalescing random walks were used to prove interface tightness, our motivation is the other way around: we aim to use interface tightness as a tool to study the dual branching-coalescing random walks. More precisely, the present paper arises out of the problem of proving convergence of rescaled branching-coalescing random walks with weak branching to a continuum object called the *Brownian net* [SS08c].

Let us first recall the graphical representation of (biased) voter models. Plot space horizontally and time vertically, and for each $i, j \in \mathbb{Z}$, at the times t of an independent Poisson point process with intensity $(1-\varepsilon)a(j-i)$, draw an arrow from (i,t) to (j,t). We call such arrows *resampling arrows*. Also, for each $i, j \in \mathbb{Z}$, at the times t of an independent Poisson process with intensity $\varepsilon a(j-i)$, draw a different type of arrow (e.g., with a different color) from (i,t) to (j,t). We call such arrows *selection arrows*.

It is well-known that a voter model can be constructed in such a way that starting from the initial state, at each time t where there is a resampling arrow from (i,t) to (j,t), the site j adopts the type of site i. To get a biased voter model one also adds the selection arrows, which are similar to resampling arrows, except that they only have an effect when the site i is of type 1.

To see the duality, we construct a system of coalescing random walks evolving backwards in time as follows: let the backward random walk at site j jump to i when it meets a resampling arrow from i to j, and let it coalesce with the random walk at site i if there is one. A system of backward branching-coalescing random walks can be obtained by moreover allowing the coalescing random walks to branch at selection arrows. That is, when a random walk at j meets a selection arrow from i to j, let it branch into two walks located at i and j, respectively. The duality goes as follows. Let A and B be two sets of integers. For the biased voter model starting from the state being 1 only on A at time 0, the set of 1's at time t has nonempty intersection with B if and only if for the backward branching-coalescing random walks starting from B at time t, there is at least one walk in A at time 0.

It is shown in [FINR04] that for coalescing nearest neighbor random walks in the space-time plane, the diffusively rescaled system converges to the so-called *Brownian web*, which, loosely speaking, is a collection of coalescing Brownian motions starting from every space-time point (see also [SSS17] for a survey on the Brownian web, Brownian net and related topics). Later, in [NRS05], this result was extended to general coalescing random walks with a finite fifth moment. This condition was then relaxed to a finite $(3+\eta)$-th moment by Belhaouari et al. [BMSV06]. It was observed in the same article (see [BMSV06, Theorem 1.2]) that to verify tightness for rescaled systems of random walks, it suffices to show that for the dual voter model starting from the heaviside state x_0 (cf. 1.4), the trajectories of the left and the right interface boundaries converge to the same Brownian motion. In the biased case, Sun and Swart [SS08c] showed that systems of branching-coalescing nearest neighbor random walks converge to the Brownian net as the branching rate ε tends to zero and space and time are diffusively rescaled with the same ε. Just as in the unbiased case, in order to extend the result of [SS08c] to non-nearest neighbor random walks, tightness can be established by showing that the interface boundaries of the dual biased voter model converge to the same Brownian motion, which in contrast to the unbiased case, has a drift. A first step in this direction is to prove that interface tightness holds in a way that is uniform as the bias parameter ε decreases to zero, which led to our Theorem 1.3. We also remark that based on the interface tightness result obtained in [Val10], Mountford and Valesin [MV16] showed that the interface of the multi-type contact process has Brownian motion as its scaling limit.

1.4 Some open problems

Before closing the introduction, we list some open problems in two directions. First, while believed to be a common phenomenon, interface tightness has so far been proved only for a few systems, including (biased) voter models and multi-type contact processes. Staying within the class of biased voter models, since there are two types of arrows, it would be meaningful to consider models with different kernels for resampling and selection arrows. It seems plausible that for such models, Theorems 1.2 and 1.3 still hold if both kernels satisfy our moment assumptions, but unfortunately our methods break down if the kernels are different. Another class of models for which the question of interface tightness remains open are voter models with weak heterozygosity selection, such as the rebellious voter model introduced in [SS08a].

Another possible direction of further research is concerned with questions related to the models studied in Theorems 1.2 and 1.3. In particular, sending the bias ε to zero and at the same time rescaling space with a factor ε and time with a factor ε^2, one would like to show that the interface converges to a Brownian motion with drift and then use this to prove convergence of the dual system of branching-coalescing random walks to the Brownian net.

The rest of the paper is devoted to proofs. We prove Theorem 1.2 in Section 2 and Theorem 1.3 in Section 3.

2 Interface tightness

2.1 Elementary observations and outline

Before moving to the proof of Theorem 1.2, we present the following elementary lemma that shows the irreducibility of non-nearest neighbor biased voter models.

Lemma 2.1 (Irreducibility of the biased voter model) *Assume that the kernel a is non-nearest neighbor (i.e., $a(k) > 0$ for some $|k| \geq 2$), irreducible and satisfies $\sum_k a(k)|k| < \infty$. Then the biased voter model $(X^\varepsilon_t)_{t\geq 0}$ on $S^{01}_{\rm int}$ is irreducible.*

Proof. As proved in [BMSV06] and also in Lemma 2.8 below, the condition $\sum_k a(k)|k| < \infty$ guarantees that the continuous-time Markov chain on $S^{01}_{\rm int}$ is well-defined and nonexplosive.

We will show that for any configuration $x \in S^{01}_{\rm int}$, 1) for each $i \in \mathbb{Z}$, there is a positive probability to reach the heaviside state $x_i := 1_{\{i,i+1,\dots\}}$ from x, and 2) there exists some $i \in \mathbb{Z}$ such that there is a positive probability to reach x from x_i. The first statement actually holds for any irreducible a. For such a, there exist $k_{\rm r}, k_{\rm l} > 0$ such that $a(k_{\rm r}), a(-k_{\rm l}) > 0$. Since $a(k_{\rm r}) > 0$ (resp. $a(-k_{\rm l}) > 0$), with positive probability the leftmost 1 (resp. the rightmost 0) can change into a 0 (resp. 1) while all other sites remain unchanged. In this way, x_i can be reached for any $i \in \mathbb{Z}$.

For the second statement, let k be such that $a(k) > 0$ for some $|k| \geq 2$. We will prove that if $k < 0$ (resp. $k > 0$), then the interface can always be expanded by 1 unit at the right (resp. left) boundary without changing the values at other sites, which implies statement 2). By symmetry, it suffices to consider the case $k < 0$. Suppose that the right boundary of X_t^ε is at site i, namely $X_t^\varepsilon(i) = 0$ and $X_t^\varepsilon(j) = 1$ for all $j > i$. Then we will construct infections to show that at a later time s, with positive probability, it may happen that $X_s^\varepsilon(i, i+1) = 00$ (similarly, it may happen that $X_s^\varepsilon(i, i+1) = 10$), while the values of other sites of X_t^ε and X_s^ε are the same. Indeed, since a is irreducible, a path π of infections from site i to $i+1$ can be constructed where moreover the path first does right jumps, and then left jumps. Recall that the value of site i is 0. Thus, the value of site $i+1$ is altered to 0 and all sites on the left of i remains unchanged. Now by applying left infections of size k, one can consecutively alter the values back to 1 for those sites on the right of $i+1$ that were infected by π. After such infections, we have $X_s^\varepsilon(i, i+1) = 00$ while other sites remain unchanged. To see that the value of site i can also be altered to 1, simply note that site $i + |k|$ can infect site i by a left infection of size k. Furthermore, we have $X_s^\varepsilon(i + |k|) = 1$ since $|k| \geq 2$. This completes the proof. ∎

The irreducibility of the biased voter model $(X_t^\varepsilon)_{t\geq 0}$ immediately implies the irreducibility of the translated process $(\overline{X}_t^\varepsilon)_{t\geq 0}$. Also note that S_{int}^{01}, and hence $\overline{S}_{\text{int}}^{01}$, is countable. Therefore, by Definition 1.1, $(X_t^\varepsilon)_{t\geq 0}$ exhibits interface tightness if and only if there exists an invariant probability measure for $(\overline{X}_t^\varepsilon)_{t\geq 0}$, where the latter implies that the interface being large at a fixed time has a small probability. More precisely, let $L : S_{\text{int}}^{01} \to \mathbb{N}$, defined as

$$L(x) := \max\{i : x(i) = 0\} - \min\{i : x(i) = 1\} + 1, \tag{2.1}$$

be the interface length. Then interface tightness is equivalent to the family of random variables $\big(L(X_t^\varepsilon)\big)_{t\geq 0}$ being tight. By definition, an *inversion* is a pair (i, j) such that $j < i$ and $x(j, i) = 10$. We let $h : S_{\text{int}}^{01} \to \mathbb{R}$ denote the function counting the number of inversions

$$h(x) := \big|\{(j, i) : j < i,\ x(j, i) = 10\}\big|. \tag{2.2}$$

It is easy to see that $h(x) \geq L(x) - 1$ for any configuration x. In the proofs of [CD95, BMV07], the number of inversions $h(x)$ plays a key role, where duality is used to show that this quantity cannot be too big. In [SS08b], h is also a key ingredient, playing a role similar to a Lyapunov function as in Foster's theorem (see e.g., [MPW17, Theorem 2.6.4]). More precisely, it is shown there that if interface tightness does not hold, then over sufficiently long time intervals, h would have to decrease on average more than it increases, contradicting the fact that $h \geq 0$.

In the rest of this section, we will adapt the method of [SS08b] to show Theorem 1.2. In the biased case, we need a different function from h, which we call the *weighted number of inversions*. In Subsection 2.2 , we will state three necessary lemmas and then prove Theorem 1.2. We then prove those three lemmas in Subsection 2.3, Subsection 2.4 and Subsection 2.5, respectively.

2.2 Proof of Theorem 1.2

Our key observation here is that the number of inversions h in (2.2) can be generalized to the biased case as follows. Let $h^\varepsilon : S_{\text{int}}^{01} \to \mathbb{R}$ be defined by

$$h^\varepsilon(x) := \sum_{i>j} (1-\varepsilon)^{\sum_{n<j} x(n)} 1_{\{x(j,i)=10\}}. \tag{2.3}$$

By numbering the 1's in x from left to right, we see that $h^\varepsilon(x)$ is a weighted number of inversions, where inversions involving the j-th 1 carry weight $(1-\varepsilon)^{j-1}$. It is clear that for every configuration $x \in S_{\text{int}}^{01}$, $h^0(x)$ agrees with the number of inversions $h(x)$ given in (2.2), and $h^\varepsilon(x) \to h^0(x)$ as $\varepsilon \downarrow 0$.

To prove Theorem 1.2, we need three lemmas that generalize Lemma 2, Lemma 3 and Proposition 4 of [SS08b] to the biased voter model. To state the first lemma that gives an expression for the action of the generator from (1.1) on h^ε, we introduce the following notation.

By definition, a *k-boundary* is a pair $(i, i+k)$ such that $x(i) \neq x(i+k)$. For $k \in \mathbb{Z}$, let $I_k : S^{01}_{\rm int} \to \mathbb{N}$ be the function counting the number of k-boundaries

$$I_k(x) := \big|\{i : x(i) \neq x(i+k)\}\big|, \tag{2.4}$$

where $|\cdot|$ denotes the cardinality of a set.

Lemma 2.2 (Generator calculations) *Under the assumption of Theorem 1.2, we have that for any $\varepsilon \in [0,1)$ and $x \in S^{01}_{\rm int}$,*

$$G^\varepsilon h^\varepsilon(x) = \sum_k a(k)\big(\tfrac{1}{2}k^2 - \varepsilon R^\varepsilon_k(x)\big) - \tfrac{1}{2}\sum_k a(k) I_k(x), \tag{2.5}$$

where the generator G^ε is defined in (1.1)*, and the term $R^\varepsilon_k(x) \geq 0$ is given by $R^\varepsilon_k(x) := 0$ for $k = -1, 0, 1$ and*

$$R^\varepsilon_k(x) := \begin{cases} \displaystyle\sum_i \sum_{n=1}^{k-1} (1-\varepsilon)^{\sum_{j<i} x(j)}(k-n) 1_{\{x(i-n,i)=01\}} & (k > 1), \\ \displaystyle\sum_i \sum_{n=1}^{|k|-1} (1-\varepsilon)^{\sum_{j<i} x(j)}(|k|-n) 1_{\{x(i,i+n)=10\}} & (k < -1). \end{cases} \tag{2.6}$$

Moreover, for $\varepsilon \in (0,1)$, we have

$$G^\varepsilon h^\varepsilon(x) \leq \tfrac{1}{2}\sum_{k<0} a(k)k^2 + \varepsilon^{-1}\sum_{k>0} a(k)k - \tfrac{1}{2}\sum_k a(k) I_k(x). \tag{2.7}$$

Remark 2.3 In the unbiased case $\varepsilon = 0$, (2.5) reduces to

$$G^0 h^0(x) = \tfrac{1}{2}\sum_k a(k)\big(k^2 - I_k(x)\big), \tag{2.8}$$

which agrees with [SS08b, Lemma 2]. Since $R^\varepsilon_k \geq 0$ for all $k \neq 0$, (2.5) shows that $G^0 h^0$ is an upper bound of $G^\varepsilon h^\varepsilon$ uniformly in ε in the sense that $G^\varepsilon h^\varepsilon(x) \leq G^0 h^0(x)$.

Remark 2.4 In a sense, the best motivation we can give on why the weighted number of inversions h^ε as in (2.3) is the "right" function to look at is formula (2.5). In the nearest-neighbor case $a(-1) = a(1) = \frac{1}{2}$ and $a(k) = 0$ for all $k \neq -1, 1$, formula (2.5) reduces to $G^\varepsilon h^\varepsilon(x) = \frac{1}{2}(1 - I_1(x))$. In particular, $G^\varepsilon h^\varepsilon(x) = -1$ if $I_1(x) = 3$, which can be used to prove (compare [SS15, Lemma 12]) that in the nearest-neighbor case, starting from an initial state with $I_1(x) = 3$, $h^\varepsilon(x)$ is the expected time before the system reaches a heaviside state. These observations originally motivated us to define h^ε as in (2.3).

We need two more lemmas. Recall that $(X^\varepsilon_t)_{t\geq 0}$ denotes the biased voter model with bias ε.

Lemma 2.5 (Nonnegative expectation) *Let the biased voter model $(X^\varepsilon_t)_{t\geq 0}$ start from a fixed configuration $X^\varepsilon_0 = x \in S^{01}_{\rm int}$. Assume either condition* (A) *or* (B) *as follows.*

(A) *$\varepsilon = 0$ and $\sum_k a(k)k^2 < \infty$.*

(B) *$\varepsilon \in (0,1)$ and $\sum_{k<0} a(k)k^2 + \sum_{k>0} a(k)k < \infty$.*

Then for any $t \geq 0$,

$$\mathbb{E}[h^\varepsilon(X_0^\varepsilon)] + \int_0^t \mathbb{E}\big[G^\varepsilon h^\varepsilon(X_s^\varepsilon)\big]\mathrm{d}s \geq 0, \tag{2.9}$$

where h^ε is the weighted number of inversions given in (2.3).

Lemma 2.6 (Interface growth) *Let $\sum_k a(k)|k| < \infty$, let a be irreducible, and let $\varepsilon \in [0,1)$. Assume that interface tightness for $(X_t^\varepsilon)_{t\geq 0}$ in the sense of Definition 1.1 does not hold on $S_{\rm int}^{01}$. Then the process started in any initial state $X_0^\varepsilon = x \in S_{\rm int}^{01}$ satisfies*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{P}\big[I_k(X_t^\varepsilon) < N\big] = 0 \qquad (k > 0,\ N < \infty), \tag{2.10}$$

where I_k is given in (2.4). The same statement holds with $S_{\rm int}^{01}$ replaced by $S_{\rm int}^{10}$.

With the lemmas above, we are ready to prove Theorem 1.2.

Proof of Theorem 1.2. By symmetry, it suffices to consider the case with state space $S_{\rm int}^{01}$. Let the biased voter model $(X_t^\varepsilon)_{t\geq 0}$ start from the heaviside state x_0 as in (1.4). Since a is irreducible, for any constant C, there exist some $k_0 \in \mathbb{Z}$ and $N \geq 1$ such that

$$C < \tfrac{1}{2} N a(k_0). \tag{2.11}$$

Let $C = \frac{1}{2}\sum_k a(k)k^2$ if a has finite second moment, and $C = \frac{1}{2}\sum_{k<0} a(k)k^2 + \varepsilon^{-1}\sum_{k>0} a(k)k$ if $\varepsilon > 0$ and the moment condition is relaxed to $\sum_{k<0} a(k)k^2 < \infty$ and $\sum_{k>0} a(k)k < \infty$. Then for all $T > 0$,

$$\begin{aligned} 0 \leq & \frac{1}{T}\int_0^T \mathbb{E}\big[G^\varepsilon h^\varepsilon(X_t^\varepsilon)\big]\mathrm{d}t \\ \leq & C - \frac{1}{2T}\sum_k a(k)\int_0^T \mathbb{E}\big[I_k(X_t^\varepsilon)\big]\mathrm{d}t \leq C - \frac{Na(k_0)}{2T}\int_0^T \mathbb{P}\big[I_{k_0}(X_t^\varepsilon) \geq N\big]\mathrm{d}t, \end{aligned} \tag{2.12}$$

where in the first inequality we used Lemma 2.5 and noted that $\mathbb{E}[h^\varepsilon(x_0)] = 0$, and in the second inequality we used Lemma 2.2, in particular, the expression (2.5) of $G^\varepsilon h^\varepsilon$ and the inequalities $R_k^\varepsilon \geq 0$ and (2.7). But on the other hand, if interface tightness did not hold, then by Lemma 2.6,

$$\lim_{T\to\infty}\Big\{C - \frac{Na(k_0)}{2T}\int_0^T \mathbb{P}\big[I_{k_0}(X_t^\varepsilon \geq N)\big]\mathrm{d}t\Big\} = C - \tfrac{1}{2}Na(k_0) < 0, \tag{2.13}$$

which contradicts (2.12). Thus interface tightness must hold for the biased voter model. ∎

2.3 Proof of Lemma 2.2

The proof is completed via a long calculation. We first change some expressions into nice forms for later calculations. Recall from (2.3) that for $\varepsilon \in [0,1)$ and $x \in S_{\rm int}^{01}$,

$$h^\varepsilon(x) = \sum_i 1_{\{x(i)=0\}} \sum_{j=-\infty}^{i-1} (1-\varepsilon)^{\sum_{n<j} x(n)} 1_{\{x(j)=1\}}. \tag{2.14}$$

Since the sum $\sum_{j=0}^{i-1}(1-\varepsilon)^j$ is $\varepsilon^{-1}\big(1-(1-\varepsilon)^i\big)$ when $\varepsilon > 0$, we can rewrite $h^\varepsilon(x)$ as

$$h^\varepsilon(x) = \begin{cases} \displaystyle\sum_{i>j}\big(1-x(i)\big)x(j) & \varepsilon = 0, \\ \displaystyle\varepsilon^{-1}\sum_i \big(1-x(i)\big)\big(1-(1-\varepsilon)^{\sum_{j<i} x(j)}\big) & \varepsilon > 0. \end{cases} \tag{2.15}$$

For each $i \in \mathbb{Z}$, define functions

$$
f_i^\varepsilon(x) := \begin{cases} \sum_{j<i} x(j) & \text{if } \varepsilon = 0, \\ \varepsilon^{-1}\big(1-(1-\varepsilon)^{\sum_{j<i} x(j)}\big) & \text{if } \varepsilon > 0, \end{cases}
\qquad g_i(x) := 1 - x(i) = 1_{\{x(i)=0\}}, \tag{2.16}
$$

and therefore

$$
h^\varepsilon = \sum_i f_i^\varepsilon g_i. \tag{2.17}
$$

We also rewrite the generator (1.1) of the biased voter model as

$$
G^\varepsilon = \sum_{k\neq 0} a(k) G_k^\varepsilon, \tag{2.18}
$$

where G_k^ε denotes the generator

$$
\begin{aligned} G_k^\varepsilon f(x) := & \sum_n 1_{\{x(n-k,n)=10\}}\big\{f(x+e_n) - f(x)\big\} \\ & + (1-\varepsilon)\sum_n 1_{\{x(n-k,n)=01\}}\big\{f(x-e_n) - f(x)\big\}, \end{aligned} \tag{2.19}
$$

which only describes infections of size k.

We start the calculations by recalling the following useful fact. Let X be a Markov process with countable state space S and generator of the form

$$
Gf(x) = \sum_y r(x,y)\big\{f(y) - f(x)\big\}, \tag{2.20}
$$

where $r(x,y)$ is the rate of jumps from configuration x to y. Then for two real functions f and g, by a direct calculation we have

$$
G(fg) = fGg + gGf + \Gamma(f,g), \tag{2.21}
$$

where

$$
\Gamma(f,g) := \sum_y r(x,y)\big\{f(y) - f(x)\big\}\big\{g(y) - g(x)\big\}, \tag{2.22}
$$

as long as all the terms involved are absolutely summable.

To find $G^\varepsilon h^\varepsilon$ for the biased voter model, applying formula (2.18) we can first calculate $G_k^\varepsilon h^\varepsilon$ and then sum over k. By (2.17) and (2.21), we have

$$
\begin{aligned} G_k^\varepsilon h^\varepsilon = & \sum_i G_k^\varepsilon(f_i^\varepsilon g_i) = \sum_i \big\{f_i^\varepsilon G_k^\varepsilon g_i + g_i G_k^\varepsilon f_i^\varepsilon - \Gamma_k^\varepsilon(f_i^\varepsilon, g_i)\big\} \\ = & \sum_i \big\{f_i^\varepsilon G_k^\varepsilon g_i + g_i G_k^\varepsilon f_i^\varepsilon\big\}, \end{aligned} \tag{2.23}
$$

where in the last step we used that $\Gamma_k^\varepsilon(f_i^\varepsilon, g_i) = 0$, since any transition either changes the value of $x(i)$, in which case f_i^ε does not change, or the transition does not change the value of $x(i)$, in which case g_i does not change.

We will prove that

$$
G_k^\varepsilon h^\varepsilon(x) = \tfrac{1}{2}\big(k^2 - I_k(x)\big) - \varepsilon R_k^\varepsilon(x). \tag{2.24}
$$

Formula (2.5) follows from this by summing over k in $\mathbb{Z}$ with weights given by a. We distinguish the calculation of $G_k^\varepsilon h^\varepsilon$ into two cases, namely $k > 0$ and $k < 0$.

Case $k > 0$. To ease notation, let us define a function $J_\varepsilon : \mathbb{N} \to \mathbb{R}$ by

$$J_\varepsilon(n) := \begin{cases} n & \text{if } \varepsilon = 0, \\ \varepsilon^{-1}\big(1 - (1-\varepsilon)^n\big) & \text{if } \varepsilon > 0, \end{cases} \tag{2.25}$$

and thus for all $\varepsilon \in [0, 1)$,

$$\begin{aligned} J_\varepsilon(n+1) - J_\varepsilon(n) &= (1-\varepsilon)^n, \\ f_i^\varepsilon(x) &= J_\varepsilon\Big(\sum_{j<i} x(j)\Big). \end{aligned} \tag{2.26}$$

Recall that $g_i(x) = 1_{\{x(i)=0\}}$. We have

$$\begin{aligned} G_k^\varepsilon f_i^\varepsilon(x) = &(1-\varepsilon)\sum_{n<i} 1_{\{x(n-k,n)=01\}}\big\{J_\varepsilon\big(\sum_{j<i} x(j) - 1\big) - J_\varepsilon\big(\sum_{j<i} x(j)\big)\big\} \\ &+ \sum_{n<i} 1_{\{x(n-k,n)=10\}}\big\{J_\varepsilon\big(\sum_{j<i} x(j) + 1\big) - J_\varepsilon\big(\sum_{j<i} x(j)\big)\big\} \\ = &-\sum_{n<i} 1_{\{x(n-k,n)=01\}}(1-\varepsilon)(1-\varepsilon)^{\sum_{j<i} x(j) - 1} \\ &+ \sum_{n<i} 1_{\{x(n-k,n)=10\}}(1-\varepsilon)^{\sum_{j<i} x(j)} \\ = &(1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n<i} \big\{1_{\{x(n-k,n)=10\}} - 1_{\{x(n-k,n)=01\}}\big\} \\ = &(1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n<i} \big\{1_{\{x(n-k)=1\}} - 1_{\{x(n)=1\}}\big\} \\ = &-(1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{k} 1_{\{x(i-n)=1\}}, \\ G_k^\varepsilon g_i(x) = &(1-\varepsilon)1_{\{x(i-k,i)=01\}} - 1_{\{x(i-k,i)=10\}} \\ = &\big(1_{\{x(i-k)=0\}} - 1_{\{x(i)=0\}}\big) - \varepsilon 1_{\{x(i-k,i)=01\}}, \end{aligned} \tag{2.27}$$

In the calculations above we used the equality

$$1_{\{x(i,j)=10\}} - 1_{\{x(i,j)=01\}} = 1_{\{x(i)=1\}} - 1_{\{x(j)=1\}} = -1_{\{x(i)=0\}} + 1_{\{x(j)=0\}}, \tag{2.28}$$

which will be used repeatedly. Substituting (2.27) into (2.23) leads to

$$\begin{aligned} G_k^\varepsilon h^\varepsilon(x) = &\sum_i f_i^\varepsilon(x)\big(1_{\{x(i-k)=0\}} - 1_{\{x(i)=0\}}\big) - \varepsilon \sum_i f_i^\varepsilon(x) 1_{\{x(i-k,i)=01\}} \\ &- \sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{k} 1_{\{x(i-n,i)=10\}}. \end{aligned} \tag{2.29}$$

Note that in general if f, g are functions such that $f_i g_i \to 0$ as $i \to \pm\infty$, then because

$$f_i g_i - f_{i-n} g_{i-n} = \big(f_i - f_{i-n}\big) g_{i-n} + f_i\big(g_i - g_{i-n}\big), \tag{2.30}$$

one has the summation by parts formula

$$\sum_i \big(f_{i+n} - f_i\big) g_i = -\sum_i f_i\big(g_i - g_{i-n}\big). \tag{2.31}$$

Applying this to $f_i(x) = f_i^\varepsilon(x)$ and $g_i(x) = 1_{\{x(i)=0\}}$, and using that $f_i^\varepsilon(x) \to 0$ as $i \to -\infty$ and $1_{\{x(i)=0\}} \to 0$ as $i \to \infty$, we have

$$-\sum_i f_i^\varepsilon(x)\big(1_{\{x(i)=0\}} - 1_{\{x(i-k)=0\}}\big) = \sum_i \big(f_{i+k}^\varepsilon(x) - f_i^\varepsilon(x)\big) 1_{\{x(i)=0\}}. \tag{2.32}$$

We moreover observe that by (2.26)

$$J_\varepsilon(n+k) - J_\varepsilon(n) = \sum_{m=0}^{k-1} \big(J_\varepsilon(n+m+1) - J_\varepsilon(n+m)\big) = \sum_{m=0}^{k-1} (1-\varepsilon)^{n+m}, \tag{2.33}$$

and hence

$$\begin{aligned} f^\varepsilon_{i+k}(x) - f^\varepsilon_i(x) &= J_\varepsilon\big(\sum_{j<i+k} x(j)\big) - J_\varepsilon\big(\sum_{j<i} x(j)\big) \\ &= \sum_{n=0}^{k-1} 1_{\{x(i+n)=1\}} (1-\varepsilon)^{\sum_{j<i+n} x(j)}. \end{aligned} \tag{2.34}$$

Inserting the identities (2.32) and (2.34) into (2.29) gives

$$\begin{aligned} G^\varepsilon_k h^\varepsilon(x) = &\sum_i \sum_{n=0}^{k-1} 1_{\{x(i,i+n)=01\}} (1-\varepsilon)^{\sum_{j<i+n} x(j)} \\ &- \sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{k} 1_{\{x(i-n,i)=10\}} \\ &- \varepsilon \sum_i f^\varepsilon_i(x) 1_{\{x(i-k,i)=01\}}. \end{aligned} \tag{2.35}$$

Before further simplifying (2.35), define counting functions $I^{10}_k, I^{01}_k : S^{01}_{\text{int}} \to \mathbb{N}$ for all $k \in \mathbb{Z}$ by

$$I^{10}_k(x) := \big|\{i : x(i,i+k) = 10\}\big| \qquad \text{and} \qquad I^{01}_k(x) := \big|\{i : x(i,i+k) = 01\}\big|. \tag{2.36}$$

Hence, by the definition (2.4) of the number of k-boundaries, we have $I_k = I^{10}_k + I^{01}_k$. Moreover, we observe that for $x \in S^{01}_{\text{int}}$, $k > 0$ and $i \in \mathbb{Z}$, along the subsequence $x(\ldots, i-k, i, i+k, \ldots)$, there is one more adjacent pair (01) than (10), and thereby for any $k > 0$,

$$I^{01}_k(x) = I^{10}_k(x) + k \qquad \text{and} \qquad I_k(x) = 2I^{10}_k(x) + k. \tag{2.37}$$

Changing the summation order, replacing i by $i-n$ in the first term in the right-hand side of (2.35), this term becomes

$$\sum_i \sum_{n=0}^{k-1} 1_{\{x(i-n,i)=01\}} (1-\varepsilon)^{\sum_{j<i} x(j)}, \tag{2.38}$$

we can rewrite (2.35) as

$$
\begin{aligned}
G_k^\varepsilon h^\varepsilon(x) = & \sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{k-1} \big\{ 1_{\{x(i-n,i)=01\}} - 1_{\{x(i-n,i)=10\}} \big\} \\
& - \sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} 1_{\{x(i-k,i)=10\}} \\
& - \varepsilon \sum_i f_i^\varepsilon(x) 1_{\{x(i-k,i)=01\}} \\
= & \sum_i \sum_{n=1}^{k-1} \big\{ 1_{\{x(i-n,i)=01\}} - 1_{\{x(i-n,i)=10\}} \big\} \\
& - \sum_i \big\{ 1 - (1-\varepsilon)^{\sum_{j<i} x(j)} \big\} \sum_{n=1}^{k-1} \big\{ 1_{\{x(i-n,i)=01\}} - 1_{\{x(i-n,i)=10\}} \big\} \\
& - \sum_i 1_{\{x(i-k,i)=10\}} + \sum_i \big\{ 1 - (1-\varepsilon)^{\sum_{j<i} x(j)} \big\} 1_{\{x(i-k,i)=10\}} \\
& - \varepsilon \sum_i f_i^\varepsilon(x) 1_{\{x(i-k,i)=01\}} \\
= & \sum_{n=1}^{k-1} \big(I_n^{01}(x) - I_n^{10}(x) \big) - I_k^{10}(x) \\
& - \varepsilon \sum_i f_i^\varepsilon(x) \sum_{n=1}^{k-1} \big\{ 1_{\{x(i-n,i)=01\}} - 1_{\{x(i-n,i)=10\}} \big\} \\
& + \varepsilon \sum_i f_i^\varepsilon(x) 1_{\{x(i-k,i)=10\}} - \varepsilon \sum_i f_i^\varepsilon(x) 1_{\{x(i-k,i)=01\}} \\
= & \sum_{n=1}^{k-1} n - \tfrac{1}{2}(I_k(x) - k) - \varepsilon \sum_i f_i^\varepsilon(x) \sum_{n=1}^{k} \big\{ 1_{\{x(i-n,i)=01\}} - 1_{\{x(i-n,i)=10\}} \big\} \\
= & \tfrac{1}{2}\big(k^2 - I_k(x)\big) - \varepsilon \sum_{n=1}^{k} \sum_i f_i^\varepsilon(x) \big(1_{\{x(i-n)=0\}} - 1_{\{x(i)=0\}} \big).
\end{aligned}
\tag{2.39}
$$

where in the third and fourth equalities we used (2.36) and (2.37), respectively, and in the last equality we interchanged the order of summation and used (2.28). We then substitute (2.32) back into the sum in the last line of (2.39), and use (2.34) to obtain

$$
\begin{aligned}
\sum_{n=1}^{k} \sum_i f_i^\varepsilon(x) \big(1_{\{x(i-n)=0\}} - 1_{\{x(i)=0\}} \big) &= \sum_{n=1}^{k} \sum_i \big(f_{i+n}^\varepsilon(x) - f_i^\varepsilon(x) \big) 1_{\{x(i)=0\}} \\
&= \sum_i \sum_{n=1}^{k} \sum_{m=0}^{n-1} (1-\varepsilon)^{\sum_{j<i+m} x(j)} 1_{\{x(i,i+m)=01\}} \\
&= \sum_i \sum_{m=0}^{k-1} \sum_{n=m+1}^{k} (1-\varepsilon)^{\sum_{j<i} x(j)} 1_{\{x(i-m,i)=01\}} \\
&= \sum_i \sum_{m=1}^{k-1} (1-\varepsilon)^{\sum_{j<i} x(j)} (k-m) 1_{\{x(i-m,i)=01\}},
\end{aligned}
\tag{2.40}
$$

which implies (2.24) for $k > 0$.

Case $k<0$**.** Similarly, we calculate

$$\begin{aligned}
G_k^\varepsilon f_i^\varepsilon(x) &= (1-\varepsilon)\sum_{n<i} 1_{\{x(n,n+|k|)=10\}}\big\{J_\varepsilon\big(\sum_{j<i} x(j)-1\big) - J_\varepsilon\big(\sum_{j<i} x(j)\big)\big\} \\
&\quad + \sum_{n<i} 1_{\{x(n,n+|k|)=01\}}\big\{J_\varepsilon\big(\sum_{j<i} x(j)+1\big) - J_\varepsilon\big(\sum_{j<i} x(j)\big)\big\} \\
&= -\sum_{n<i} 1_{\{x(n,n+|k|)=10\}}(1-\varepsilon)(1-\varepsilon)^{\sum_{j<i} x(j)-1} \\
&\quad + \sum_{n<i} 1_{\{x(n,n+|k|)=01\}}(1-\varepsilon)^{\sum_{j<i} x(j)} \\
&= (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n<i} \big\{1_{\{x(n,n+|k|)=01\}} - 1_{\{x(n,n+|k|)=10\}}\big\} \\
&= (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=0}^{|k|-1} 1_{\{x(i+n)=1\}}, \\
G_k^\varepsilon g_i(x) &= (1-\varepsilon)1_{\{x(i,i+|k|)=10\}} - 1_{\{x(i,i+|k|)=01\}} \\
&= \big(1_{\{x(i+|k|)=0\}} - 1_{\{x(i)=0\}}\big) - \varepsilon 1_{\{x(i,i+|k|)=10\}},
\end{aligned} \tag{2.41}$$

which gives

$$\begin{aligned}
G_k^\varepsilon h^\varepsilon(x) &= \sum_i f_i^\varepsilon(x)\big(1_{\{x(i+|k|)=0\}} - 1_{\{x(i)=0\}}\big) - \varepsilon\sum_i f_i^\varepsilon(x)1_{\{x(i,i+|k|)=10\}} \\
&\quad + \sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{|k|-1} 1_{\{x(i,i+n)=01\}}.
\end{aligned} \tag{2.42}$$

Since by summation by parts,

$$\begin{aligned}
\sum_i f_i^\varepsilon(x)\big(1_{\{x(i+|k|)=0\}} - 1_{\{x(i)=0\}}\big) &= -\sum_i \big(f_i^\varepsilon(x) - f_{i-|k|}^\varepsilon(x)\big)1_{x(i)=0} \\
&= -\sum_i \sum_{n=1}^{|k|} (1-\varepsilon)^{\sum_{j<i-n} x(j)} 1_{\{x(i-n,i)=10\}} \\
&= -\sum_i (1-\varepsilon)^{\sum_{j<i} x(j)} \sum_{n=1}^{|k|} 1_{\{x(i,i+n)=10\}},
\end{aligned} \tag{2.43}$$

combining the first and last terms on the right-hand side of (2.42), we can rewrite

$$\begin{aligned}
G_k^\varepsilon h^\varepsilon(x) &= \sum_i \sum_{n=1}^{|k|-1} \big\{1_{\{x(i,i+n)=01\}} - 1_{\{x(i,i+n)=10\}}\big\} \\
&\quad - \sum_i \big\{1-(1-\varepsilon)^{\sum_{j<i} x(j)}\big\} \sum_{n=1}^{|k|-1} \big\{1_{\{x(i,i+n)=01\}} - 1_{\{x(i,i+n)=10\}}\big\} \\
&\quad - \sum_i 1_{\{x(i,i+|k|)=10\}} + \sum_i \big\{1-(1-\varepsilon)^{\sum_{j<i} x(j)}\big\} 1_{\{x(i,i+|k|)=10\}} \\
&\quad - \varepsilon\sum_i f_i^\varepsilon(x) 1_{\{x(i,i+|k|)=10\}} \\
&= \sum_{n=1}^{|k|-1} \big(I_n^{01}(x) - I_n^{10}(x)\big) - I_{|k|}^{10}(x) \\
&\quad - \varepsilon\sum_i f_i^\varepsilon(x) \sum_{n=1}^{|k|-1} \big(1_{\{x(i)=0\}} - 1_{\{x(i+n)=0\}}\big), \\
&= \tfrac{1}{2}\big(k^2 - I_{|k|}(x)\big) - \varepsilon R_k^\varepsilon(x),
\end{aligned} \tag{2.44}$$

where in the last equality we have used the summation by parts formula (2.31) and then (2.34) as follows

$$
\begin{aligned}
\sum_{n=1}^{|k|-1}\sum_i f_i^\varepsilon(x)\big(1_{\{x(i)=0\}}-1_{\{x(i+n)=0\}}\big) &= \sum_{n=1}^{|k|-1}\sum_i \big(f_i^\varepsilon(x)-f_{i-n}^\varepsilon(x)\big)1_{\{x(i)=0\}}\\
&=\sum_i\sum_{n=1}^{|k|-1}\sum_{m=1}^{n}(1-\varepsilon)^{\sum_{j<i-m}x(j)}1_{\{x(i-m,i)=01\}}\\
&=\sum_i\sum_{m=1}^{|k|-1}\sum_{n=m}^{|k|-1}(1-\varepsilon)^{\sum_{j<i}x(j)}1_{\{x(i,i+m)=01\}}\\
&=\sum_i\sum_{m=1}^{|k|-1}(1-\varepsilon)^{\sum_{j<i}x(j)}(|k|-m)1_{\{x(i,i+m)=01\}}.
\end{aligned}
\tag{2.45}
$$

Since $I_{-k}(x)=I_k(x)$ according to the definition (2.4), by (2.44), we see that (2.24) holds also for $k<0$.

In order to obtain the inequality (2.7) when $\varepsilon\in(0,1)$, let $i_0:=\inf\{i\in\mathbb{Z}:x(i)=1\}$ and inductively let $i_n:=\inf\{i>i_{n-1}\in\mathbb{Z}:x(i)=1\}$, i.e., $i_0,i_1,\dots$ are the positions of the first, second etc. 1, coming from the left. Thus, by counting from the left to right, for $k>0$,

$$
\begin{aligned}
R_k^\varepsilon(x) &= \sum_{n=0}^{\infty}(1-\varepsilon)^n\sum_{m=1}^{k-1}(k-m)1_{\{x(i_n-m)=0\}}\\
&=\sum_{n=0}^{\infty}(1-\varepsilon)^n\Big(\sum_{m=1}^{k-1}(k-m)-\sum_{m=1}^{k-1}(k-m)1_{\{x(i_n-m)=1\}}\Big)\\
&\geq\sum_{n=0}^{\infty}(1-\varepsilon)^n\Big(\tfrac12 k(k-1)-kn\Big)\\
&=\tfrac12\varepsilon^{-1}k(k-1)-\varepsilon^{-2}(1-\varepsilon)k,
\end{aligned}
\tag{2.46}
$$

where in the inequality we bounded $\sum_{m=1}^{k-1}(k-m)1_{\{x(i_n-m)=1\}}$ by $k\sum_{m=1}^{k-1}1_{\{x(i_n-m)=1\}}$ and then used that there are at most n 1's on the left of site i_n, and in the last equality we used the identity $\sum_{n=0}^{\infty}(1-\varepsilon)^n n=\varepsilon^{-2}(1-\varepsilon)$. Inserting (2.46) into (2.24), we obtain, for $\varepsilon>0$,

$$
G_k^\varepsilon h^\varepsilon(x)\leq\tfrac12\big(k^2-I_k(x)\big)-\big(\tfrac12k^2-\tfrac12k-\varepsilon^{-1}k+k\big)\leq\varepsilon^{-1}k-\tfrac12 I_k(x)\quad(k>0).\tag{2.47}
$$

Using this and combining it for $k<0$ with the more elementary estimate $R_k^\varepsilon(x)\geq0$ in (2.24), and summing over k, we then obtain (2.7).

2.4 Proof of Lemma 2.5

Though Lemma 2.5 under condition (A) is exactly Lemma 3 of [SS08b], we cannot follow the proof there to show our result under condition (B). More precisely, the estimate (3.24) in [SS08b] cannot be used in case (B) due to the loss of finite second moment. Instead, our proof uses different estimates, see Lemmas 2.7 and 2.8 below, which turn out to work for the lemma under condition (A) as well.

Let us recall from (2.36) that $I_k^{10}(x)=\big|\{i:x(i,i+k)=10\}\big|$.

Lemma 2.7 (Bound on number of inversions) *Let $x\in S_{\rm int}^{01}$ and $I_n^{10}(x)$ be as in (2.4). Then for any $0\leq n<m$,*

$$
|I_m^{10}(x)-I_n^{10}(x)|\leq(m-n)I_1^{10}(x).\tag{2.48}
$$

Proof. Suppose the interface of x consists of K blocks of consecutive 1's and K blocks of consecutive 0's as follows.

$$\cdots 0000000000\underbrace{1111\cdots 1111}_{\text{1st 1 block}}\overbrace{0000\cdots 0000}^{\text{1st 0 block}}\cdots\cdots\cdots\underbrace{1111\cdots 1111}_{K\text{-th 1 block}}\overbrace{0000\cdots 0000}^{K\text{-th 0 block}}1111111111\cdots$$

It is straightforward to see that $I_1^{10}(x) = K$. Suppose that the k-th block of consecutive 1's is from site i_k to site j_k. Then

$$I_n^{10}(x) = \sum_{k=1}^{K}\sum_{s=i_k}^{j_k} 1_{\{x(s+n)=0\}}, \tag{2.49}$$

and therefore for $0 < n < m$,

$$\begin{aligned}\left|I_m^{10}(x) - I_n^{10}(x)\right| &\le \sum_{k=1}^{K}\Big|\sum_{s=i_k}^{j_k} 1_{\{x(m+s)=0\}} - \sum_{s=i_k}^{j_k} 1_{\{x(n+s)=0\}}\Big| \\ &= \sum_{k=1}^{K}\Big|\sum_{s=n+1}^{m} 1_{\{x(j_k+s)=0\}} - \sum_{s=n}^{m-1} 1_{\{x(i_k+s)=0\}}\Big| \\ &\le \sum_{k=1}^{K}\sum_{s=n+1}^{m}\left|1_{\{x(j_k+s)=0\}} - 1_{\{x(i_k+s-1)=0\}}\right| \\ &\le \sum_{k=1}^{K}(m-n) = (m-n)I_1^{10}(x).\end{aligned} \tag{2.50}$$

In particular, when $n = 1 < m$, (2.50) implies that

$$|I_m^{10}(x) - I_1^{10}(x)| \le (m-1)I_1^{10}(x), \tag{2.51}$$

which results in

$$I_m^{10}(x) \le mI_1^{10}(x). \tag{2.52}$$

The last inequality is nothing but (2.48) for the case of $n = 0 < m$, and thus the proof is complete. ∎

The following lemma bounds uniformly the expected number of 1-boundaries I_1 from (2.4).

Lemma 2.8 (Bound on 1-boundaries) *Let $\sum_k a(k)|k| < \infty$. Let $(X_t^\varepsilon)_{t\ge 0}$ be a biased voter model starting from a fixed configuration $x \in S_{\rm int}^{01}$. Then*

$$\sup_{\varepsilon\in[0,1)} \mathbb{E}\big[I_1(X_t^\varepsilon)\big] \le I_1(x)e^{Ct}, \tag{2.53}$$

where $C := 2\sum_{k\neq 0} a(k)|k-1|$.

Remark 2.9 Below in Lemma 3.2, we also give a bound on the expected number of 1-boundaries under the invariant law. Although the statements are similar, the proofs of Lemmas 2.8 and 3.2 are completely different.

Proof. We will couple the process $\big(I_1(X_t^\varepsilon)\big)_{t\ge 0}$ to a branching process $(Z_t)_{t\ge 0}$ in such a way that $I_1(X_t^\varepsilon) \le Z_t$ for all $t \ge 0$. The left-hand side of (2.53) can then be uniformly bounded from above by the expectation of Z_t, which in turn can be upper bounded by the right-hand side of (2.53). To see the coupling, note that whenever an infection of the biased voter model increases the number of 1-boundaries, it must jump across at least one 1-boundary and end at least a distance two from this 1-boundary. For each 1-boundary the rate of infections that cross it in this way is $\sum_{k\neq 0} a(k)|k-1|$. Since a single infection increases the number of 1-boundaries at most by 2, we can a.s. bound $I(X^\varepsilon_t)$ from above by a branching process $(Z_t)_{t\geq 0}$ started in $Z_0 = I_1(x)$, where each particle produces two offspring with rate $\sum_{k\neq 0} a(k)|k-1|$, leading to the bound (2.53). ∎

Proof of Lemma 2.5. By standard theory, for a bounded function f with bounded $G^\varepsilon f$,

$$M_t(f) := f(X^\varepsilon_t) - \int_0^t G^\varepsilon f(X^\varepsilon_s)\mathrm{d}s \qquad (t\geq 0) \tag{2.54}$$

is a martingale. Although the weighted number of inversions h^ε is not bounded from above, with a bit extra work we can see that (2.9) still holds by a truncation-approximation argument as in [SS08b]. Let $(X^{\varepsilon,K}_t)_{t\geq 0}$ be the process with the same initial state $X^{\varepsilon,K}_0 = X^\varepsilon_0$ and truncated kernel $a^K(k) := 1_{\{|k|\leq K\}}a(k)$ $(k\in\mathbb{Z})$. Recall that $L(x)$, defined in (2.1), denotes the interface length of a configuration $x\in S^{01}_{\mathrm{int}}$. Define stopping times

$$\tau_{K,N} := \inf\{t\geq 0 : L(X^{\varepsilon,K}_t) > N\} \quad \text{and} \quad \tau_N := \inf\{t\geq 0 : L(X^\varepsilon_t) > N\}. \tag{2.55}$$

Let $G^{\varepsilon,K}$ denote the generator from (1.1) with the kernel a replaced by a^K. For fixed K and N, since $L(X^{\varepsilon,K}_t)$ is bounded by $K+N$ for all $t\leq\tau_{K,N}$, and $h^\varepsilon(x)\leq h^0(x)$ where the latter is further bounded by $L(x)^2$, the interface length squared, we conclude that

$$M^{K,N}_t := h^\varepsilon(X^{\varepsilon,K}_{t\wedge\tau_{K,N}}) - \int_0^{t\wedge\tau_{K,N}} G^{\varepsilon,K}h^\varepsilon(X^{\varepsilon,K}_s)\mathrm{d}s \quad (t\geq 0) \tag{2.56}$$

is a martingale, which yields

$$0\leq \mathbb{E}[h^\varepsilon(X^{\varepsilon,K}_{t\wedge\tau_{K,N}})] = \mathbb{E}[h^\varepsilon(X^{\varepsilon,K}_0)] + \mathbb{E}\Big[\int_0^{t\wedge\tau_{K,N}} G^{\varepsilon,K}h^\varepsilon(X^{\varepsilon,K}_s)\mathrm{d}s\Big]. \tag{2.57}$$

We will take limits as $N,K\to\infty$. Since $X^{\varepsilon,K}_0 = X^\varepsilon_0$, the lemma would follow once we show

$$\lim_{N\to\infty}\lim_{K\to\infty}\mathbb{E}\Big[\int_0^t 1_{\{s<\tau_{K,N}\}}G^{\varepsilon,K}h^\varepsilon(X^{\varepsilon,K}_s)\mathrm{d}s\Big] = \mathbb{E}\Big[\int_0^t G^\varepsilon h^\varepsilon(X^\varepsilon_s)\mathrm{d}s\Big]. \tag{2.58}$$

First observe that we can couple X^ε and $X^{\varepsilon,K}$ such that there exists a random K_0 so that if $K>K_0$, then $\tau_{K,N}=\tau_N$ and $X^{\varepsilon,K}_t = X^\varepsilon_t$ for all $t\leq\tau_N$. In particular, almost surely,

$$\lim_{K\to\infty}\tau_{K,N}=\tau_N \quad \text{and} \quad \lim_{K\to\infty} X^{\varepsilon,K}(s)1_{\{s<\tau_{K,N}\}} = X^\varepsilon(s)1_{\{s<\tau_N\}}. \tag{2.59}$$

Next we note that almost surely, $\lim_{N\to\infty}\tau_N=\infty$. Indeed, if

$$a_{\mathrm{s}}(k) := \tfrac{1}{2}\big(a(k)+a(-k)\big) \tag{2.60}$$

denotes the symmetrization of $a(\cdot)$, then we can further couple X^ε and $X^{\varepsilon,K}$ with a uni-directional random walk S with increment rate $q(n) := \sum_{k=n}^\infty 2a_{\mathrm{s}}(k)$ $(n\geq 1)$ and $S_0 = L(X^\varepsilon_0)$, such that $L(X^\varepsilon_t)\leq S_t$ and $L(X^{\varepsilon,K}_t)\leq S_t$ for all $t\geq 0$. It is then not hard to see that $\lim_{N\to\infty}\tau_N=\infty$ almost surely.

Now recall that, similar to the expression (2.5) for $G^\varepsilon h^\varepsilon(x)$, we have

$$G^{\varepsilon,K}h^\varepsilon(X^{\varepsilon,K}_s)1_{\{s<\tau_{K,N}\}} = \sum_k 1_{\{s<\tau_{K,N}\}}1_{\{|k|\leq K\}}a(k)\big(\tfrac{1}{2}k^2 - \varepsilon R^\varepsilon_k(X^{\varepsilon,K}_s) - \tfrac{1}{2}I_k(X^{\varepsilon,K}_s)\big). \tag{2.61}$$

Since $\tau_{K,N} \to \tau_N$, $\tau_N \to \infty$, and $X^{\varepsilon,K}(s) \to X^\varepsilon(s)$, we see that for each $s \in [0,t]$, $k \in \mathbb{Z}$, and almost surely, as first $K \to \infty$ and then $N \to \infty$, the summand above converges to

$$a(k)\big(\tfrac{1}{2}k^2 - \varepsilon R_k^\varepsilon(X_s^\varepsilon)\big) - \tfrac{1}{2}a(k)I_k(X_s^\varepsilon), \tag{2.62}$$

the sum of which over k gives $G^\varepsilon h^\varepsilon(X_s^\varepsilon)$. We will extend this pointwise convergence to the convergence of their integral with respect to $\mathbb{E}\big[\int_0^t \mathrm{d}s \sum_k \cdot\big]$ in (2.58).

We first treat the contribution from $I_k(X_s^{\varepsilon,K})$ in (2.61). Note that by (2.37) and Lemma 2.7, we have

$$I_k(X_s^{\varepsilon,K}) = I_k^{10}(X_s^{\varepsilon,K}) + I_k^{01}(X_s^{\varepsilon,K}) \le |k|\big(1 + 2I_1^{10}(X_s^{\varepsilon,K})\big). \tag{2.63}$$

Recall that there exists a random K_0 such that if $K \ge K_0$, then $X_s^{\varepsilon,K} = X_s^\varepsilon$ for any $s \le \tau_{K,N}$. On the event $K_0 > K$, we have $I_1^{10}(X_s^{\varepsilon,K}) \le L(X_s^{\varepsilon,K}) + 1 \le N+1$ for all $s < \tau_{K,N}$. Therefore

$$\mathbb{E}\Big[1_{\{K_0>K\}} \int_0^{t\wedge\tau_{K,N}} \sum_{|k|\le K} a(k)I_k(X_s^{\varepsilon,K})\mathrm{d}s\Big] \le P\big(K_0 > K\big) \sum_k a(k)|k| \int_0^t \big(3+2N\big)\mathrm{d}s \tag{2.64}$$

which tends to zero in the limit of first $K \to \infty$ and then $N \to \infty$.

On the event $K_0 \le K$, because $X_s^{\varepsilon,K} = X_s^\varepsilon$ for $s < \tau_N$ and $\tau_{K,N} = \tau_N$, we have

$$1_{\{K_0\le K\}}1_{\{s<\tau_{K,N}\}}1_{\{|k|\le K\}}a(k)I_k(X_s^{\varepsilon,K}) \le a(k)|k|\big(1 + 2I_1^{10}(X_s^\varepsilon)\big), \tag{2.65}$$

where the right hand side is integrable with respect to $\mathbb{E}\big[\int_0^t \mathrm{d}s \sum_k \cdot\big]$ by Lemma 2.8, and the left hand side converges pointwise to $a(k)I_k(X_s^\varepsilon)$ as first $K \to \infty$ and then $N \to \infty$. Therefore by dominated convergence, together with (2.64), we obtain

$$\lim_{N\to\infty}\lim_{K\to\infty} \mathbb{E}\Big[\int_0^t 1_{\{s<\tau_{K,N}\}}\tfrac{1}{2}\sum_{|k|\le K} a(k)I_k(X_s^{\varepsilon,K})\mathrm{d}s\Big] = \mathbb{E}\Big[\int_0^t \tfrac{1}{2}\sum_k a(k)I_k(X_s^\varepsilon)\mathrm{d}s\Big]. \tag{2.66}$$

To treat the contribution from $\frac{1}{2}k^2 - \varepsilon R_k^\varepsilon(X_s^{\varepsilon,K})$, we note that by the expression (2.6) for $R_k^\varepsilon(x)$, it is easy to see that

$$\varepsilon R_k^\varepsilon(x) \le \varepsilon \sum_{i=1}^\infty (1-\varepsilon)^{i-1} \sum_{n=1}^{|k|-1} n = \tfrac{1}{2}|k|(|k|-1), \quad \text{hence} \quad \tfrac{1}{2}k^2 - \varepsilon R_k^\varepsilon(X_s^{\varepsilon,K}) \ge \tfrac{1}{2}|k|. \tag{2.67}$$

On the other hand, by the lower bound (2.46) on $R_k^\varepsilon(x)$ when $k > 0$ and the fact that $R_k^\varepsilon(x) \ge 0$ when $k < 0$, we have

$$\tfrac{1}{2}k^2 - \varepsilon R_k^\varepsilon(X_s^{\varepsilon,K}) \le 1_{\{k<0\}}\tfrac{1}{2}a(k)k^2 + 1_{\{k>0\}}\varepsilon^{-1}a(k)k, \tag{2.68}$$

which is integrable with respect to $\mathbb{E}\big[\int_0^t \mathrm{d}s \sum_k \cdot\big]$ by either condition (A) or (B), while the left hand side converges pointwise to $\frac{1}{2}k^2 - \varepsilon R_k^\varepsilon(X_s^\varepsilon)$ as first $K \to \infty$ and then $N \to \infty$. Dominated convergence theorem can then be applied, which together with (2.66), implies (2.58). ∎

Remark 2.10 If we assume the kernel $a(\cdot)$ has finite third moment, then we can prove that the process $M_t(h^\varepsilon) = h^\varepsilon(X_t^\varepsilon) - \int_0^t G^\varepsilon h^\varepsilon(X_s^\varepsilon)\mathrm{d}s$ is a martingale. To prove this, we only need to check the uniform integrability of $\big(h^\varepsilon(X_{t\wedge\tau_{K,N}}^{\varepsilon,K})\big)$ in K and N, because then we can take the limit on the left-hand side of (2.57) as well, and the equality in (2.57) remains valid as $K \to \infty$ and then $N \to \infty$. Recall that the uni-directional random walk S has increment rate $q(n) = \sum_{k=n}^\infty 2a_s(n)$ whose second moment is now finite since

$$\mathbb{E}\big[(S_t - S_0)^2\big] \le t\sum_n q(n)n^2 \le t\sum_k a(k)|k|^3 < \infty. \tag{2.69}$$

The uniform integrability thus follows from the fact that

$$h^\varepsilon(X^{\varepsilon,K}_{t\wedge\tau_{K,N}}) \le h^0(X^{\varepsilon,K}_{t\wedge\tau_N}) \le L(X^{\varepsilon,K}_{t\wedge\tau_{K,N}})^2 \le S_t^2 \quad \text{a.s.,} \tag{2.70}$$

where in the third inequality we used $L(X^{\varepsilon,K}_{t\wedge\tau_N}) \le S_{t\wedge\tau_N} \le S_t$.

In particular, for the unbiased voter model $(X^0_t)_{t\ge0}$, the process $(M_t(h^0))_{t\ge0}$ is a martingale. We claim, however, if we replace h^0 by h_M, the number of inversions within distance M, formally defined by

$$h_M(x) := \big|\{(i,j) : 0 < j-i \le M,\ x(i,j) = 10\}\big|, \tag{2.71}$$

then a finite second moment assumption would suffice to imply that $M_t(h_M)$ is a martingale, and therefore

$$\mathbb{E}\big[h_M(X^0_t)\big] - \mathbb{E}\big[h_M(X^0_0)\big] = \mathbb{E}\Big[\int_0^t G^0 h_M(X^0_s)\mathrm{d}s\Big]. \tag{2.72}$$

Indeed, since the inversion pairs must be inside the interface and each particle in the interface contributes to at most $2M$ pairs of inversions, for any $x \in S^{01}_{\rm int}$ we have

$$h_M(x) \le 2ML(x). \tag{2.73}$$

Thus the uniform integrability of $\big(h_M(X^{0,K}_{t\wedge\tau_{K,N}})\big)$ follows from

$$h_M(X^{0,K}_{t\wedge\tau_{K,N}}) \le 2ML(X^{0,K}_{t\wedge\tau_{K,N}}) \le 2MS_t \quad \text{a.s.} \qquad \text{and} \qquad \mathbb{E}[S_t] < \infty, \tag{2.74}$$

which only requires a to have finite second moment. Therefore, in order to show that $M_t(h_M)$ is a martingale, it remains to check the uniform integrability of $\int_0^{t\wedge\tau_{K,N}} G^{0,K} h_M(X^{0,K}_s)\mathrm{d}s$. Using the expression of $G^{0,K}h_M$ in (3.24) below, one only needs to show the uniform integrability of

$$\int_0^{t\wedge\tau_{K,N}} \Bigg(\sum_{n=1}^\infty A^K(n) I^{10}_{M+n}(X^{0,K}_s) - \sum_{n=1}^\infty A^K(n) I^{10}_{M-1-n}(X^{0,K}_s)\Bigg)\,\mathrm{d}s \qquad (K,N\ge1) \tag{2.75}$$

where $A^K(n) = \sum_{k=n}^\infty \big(a^K(k) + a^K(-k)\big)$. Estimating $I^{10}_{M+n} \le I_{M+n} \le (M+n)I_1$ and $I^{10}_{M-1-n} \le I_{M-1-n} \le |M-1-n|I_1$ and using Lemma 2.8, one gets that the first moment of $A(\cdot)$, or equivalently, the second moment of $a(\cdot)$, being finite guarantees the uniform integrability of the terms in (2.75). Thus the equality (2.72) has been proved, which we state as the following lemma. This result will be used in the proof of Proposition 3.7 later on.

Lemma 2.11 *Let the voter model $(X^0_t)_{t\ge0}$ start from a fixed configuration $X^0_0 = x \in S^{01}_{\rm int}$, and let h_M denote the number of inversions within distance M as in* (2.71)*. Assume that $\sum_k a(k)k^2 < \infty$. Then for any $t \ge 0$,*

$$\mathbb{E}\big[h_M(X^0_t)\big] - \mathbb{E}\big[h_M(X^0_0)\big] = \mathbb{E}\Big[\int_0^t G^0 h_M(X^0_s)\mathrm{d}s\Big]. \tag{2.76}$$

2.5 Proof of Lemma 2.6

We fix $\varepsilon \in [0,1)$ throughout the proof. We only state the proof for $S^{01}_{\rm int}$, the proof for $S^{10}_{\rm int}$ being the same. We start by proving the statement for $k=1$. Let $(X^\varepsilon_t)_{t\ge0}$ be started in an initial state $X^\varepsilon_0 = x \in S^{01}_{\rm int}$ and consider the "boundary process" $(Y_t)_{t\ge0}$ defined as

$$Y_t(i) := X^\varepsilon_t(i+1) - X^\varepsilon_t(i) \qquad (i\in\mathbb{Z},\ t\ge0). \tag{2.77}$$

The assumption $\sum_k a(k)|k| < \infty$ guarantees that a.s. $X^\varepsilon_t \in S^{01}_{\rm int}$ for all $t \ge 0$ and hence $(Y_t)_{t\ge0}$ is a Markov process in the space of all configurations $y \in \{-1,0,1\}^{\mathbb{Z}}$ such that $\sum_i |y(i)|$ is odd (and finite) and $\sum_{i:i\leq j} y(i) \in \{0,1\}$ for all $j \in \mathbb{Z}$. For any such configuration, we set $|y| := \sum_i |y(i)|$. Then $|Y_t| = I_1(X^\varepsilon_t)$ $(t \geq 0)$.

In the special case that $\varepsilon = 0$, the process $(Z_t)_{t\geq 0}$ defined as $Z_t(i) := |Y_t(i)|$ is also a Markov process, and in fact a cancellative spin system. In this case, we can apply [SS08a, Proposition 13] to conclude that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t\, \mathbb{P}\big[|Y_t| \leq N\big] = 0 \qquad (N < \infty). \tag{2.78}$$

We describe this in words by saying that for each N, the process spends a zero fraction of time in states y with $|y| \leq N$. In the biased case $\varepsilon > 0$, the process $(Z_t)_{t\geq 0}$ is no longer a Markov process, but we claim that the proof of [SS08a, Proposition 13] can easily be adapted to show that (2.78) still holds. To demonstrate this, we go through the main steps of that proof and show how to adapt them to our process $(Y_t)_{t\geq 0}$.

The main ingredient in the proof of [SS08a, Proposition 13] is formula (3.54) of that paper, which for our process must be reformulated as

$$\inf\big\{\mathbb{P}^y[|Y_t| = n] : |y| = n+2,\ y(i) \neq 0 \neq y(j) \text{ for some } i \neq j,\ |i-j| \leq L\big\} > 0 \tag{2.79}$$

for all $t > 0$, $L \geq 1$, and $n = 1, 3, 5, \ldots$. Let us call a site $i \in \mathbb{Z}$ such that $Y_t(i) \neq 0$ a boundary of X^ε_t. Then (2.79) says that if X^ε_t contains $n+2$ boundaries of which two are at distance $\leq L$ of each other, then there is a uniformly positive probability that after time t the number of boundaries has decreased by 2.

The assumption that interface tightness for X^ε_t does not hold on $S^{01}_{\rm int}$ implies that (2.78) holds for $N = 1$. The proof of (2.78) now proceeds by induction on N. Imagine that (2.78) holds for N. Then it can be shown that (2.79) implies that for each $L \geq 1$, the process spends a zero fraction of time in states y with $|y| = N+2$ which contain two boundaries at distance $\leq L$ of each other. Now imagine that (2.78) does not hold for $N+2$. Then for each $L \geq 1$, the process must spend a positive fraction of time in states y with $|y| = N+2$ but which do not contain two boundaries at distance $\leq L$ of each other.

If L is large, then each boundary evolves for a long time as a process started in a heaviside initial state, either of type 01 or of type 10, without feeling the other boundaries, which are far away. By our assumption that interface tightness on $S^{01}_{\rm int}$ does not hold, the boundaries of type 01 are unstable in the sense that they will soon split into three or more boundaries and on sufficient long time scales spend most of their time being three or more boundaries, rather than one. With some care, it can be shown that this implies that the process spends a zero fraction of time in states y with $|y| = N+2$, completing the induction step. This argument is written down more carefully in [SS08a]. Formula (3.65) of that paper has to be slightly modified in our situation since we only know that the boundaries of type 01 are unstable. So instead of producing at least $3(N+2)$ boundaries with probability at least $(1-2\varepsilon)^{N+2}$, in our case, we produce at least $3(N+3)/2 + (N+1)/2$ boundaries with probability at least $(1-2\varepsilon)^{(N+3)/2}$, since of the $N+2$ boundaries there are $(N+3)/2$ of type 01 and $(N+1)/2$ of type 10.

Translated back to the biased voter model $(X^\varepsilon_t)_{t\geq 0}$ started in an initial state $X^\varepsilon_0 = x \in S^{01}_{\rm int}$, formula (2.78) says that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{P}[I_1(X^\varepsilon_t) < N]\mathrm{d}t = 0 \qquad (N < \infty). \tag{2.80}$$

To deduce (2.10) from (2.80), it suffices to prove that for any $k, M \geq 1$ and $s > 0$,

$$\lim_{N\to\infty} \inf_{X^\varepsilon_0 : I_1(X^\varepsilon_0) \geq N} \mathbb{P}[I_k(X^\varepsilon_s) < M] = 0. \tag{2.81}$$

For if (2.81) holds, then for any $s, \delta > 0$, there exists N large enough such that the infimum in (2.81) is less than δ, which, letting the process in (2.80) evolve for some extra time s implies that

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{P}[I_k(X^\varepsilon_{t+s}) < M] < 2\delta \quad (M \geq 1). \tag{2.82}$$

As δ is arbitrary, (2.10) is obtained.

It remains to show (2.79) and (2.81). Both of them can be proved by directly constructing specific paths with positive probabilities, by constructions very similar to those in the proof of [SS08b, Prop. 4].

3 Continuity of the invariant law

3.1 Proof outline

As a positive recurrent Markov chain on the countable state space $\overline{S}^{01}_{\rm int}$, the biased voter model modulo translations has a unique invariant law $\overline{\nu}_\varepsilon$. This section is devoted to showing Theorem 1.3, namely the weak convergence $\overline{\nu}_\varepsilon \Rightarrow \overline{\nu}_0$ with respect to the discrete topology on $\overline{S}^{01}_{\rm int}$.

Recall that $\overline{S}^{01}_{\rm int}$ is the set of equivalence classes of elements of $S^{01}_{\rm int}$ that are equal up to translations. It will be convenient to choose a representative from each equivalence class by shifting the leftmost one to the origin. Since each equivalence class $\overline{x} \in \overline{S}^{01}_{\rm int}$ contains a unique element x in the set

$$\hat{S}^{01}_{\rm int} := \{x \in S^{01}_{\rm int} : x(i) = 0 \text{ for all } i < 0 \text{ and } x(0) = 1\}, \tag{3.1}$$

we can identify $\overline{S}^{01}_{\rm int}$ with $\hat{S}^{01}_{\rm int}$. Under this identification, $\overline{\nu}_\varepsilon$ on $\overline{S}^{01}_{\rm int}$ uniquely determines a probability measure ν_ε on $\{0,1\}^{\mathbb{Z}}$ that is supported on $\hat{S}^{01}_{\rm int}$. We let X^ε_∞ denote a random variable with law ν_ε.

If we could show tightness for the length $L(X^\varepsilon_\infty)$ of the interface (as defined in (2.1)), then it would be relatively straightforward to prove weak continuity of the map $\varepsilon \mapsto \nu_\varepsilon$ with respect to the discrete topology on $\hat{S}^{01}_{\rm int}$. Unfortunately, our proof of interface tightness gives us very little direct control on $L(X^\varepsilon_\infty)$. However, our methods can be used to give a uniform upper bound for the expected number of boundaries $I_1(X^\varepsilon_\infty)$ (as defined in (2.4)) and this is sufficient to prove weak continuity of the map $\varepsilon \mapsto \nu_\varepsilon$ with respect to the product topology on $\hat{S}^{01}_{\rm int}$.

We do not know if the map $\varepsilon \mapsto \nu_\varepsilon$ is continuous with respect to the discrete topology on $\hat{S}^{01}_{\rm int}$, but we can establish continuity at $\varepsilon = 0$ by a rather subtle argument. Let $\varepsilon_n \downarrow 0$ and assume that $\nu_{\varepsilon_n} \Rightarrow \nu_0$ with respect to the product topology on $\hat{S}^{01}_{\rm int}$ but not with respect to the discrete topology. Then, with positive probability, $X^{\varepsilon_n}_\infty$ must contain boundaries that "walk away to infinity" as $\varepsilon_n \downarrow 0$. We can rule out this scenario by proving that it would violate the equilibrium equation $\mathbb{E}[G^0 h^0(X^0_\infty)] = 0$, which determines the expected value of the weighted number of k-boundaries in the equilibrium voter model interface (cf. 3.2). In this way, we can prove Theorem 1.3 which indirectly also establishes tightness for the length $L(X^{\varepsilon_n}_\infty)$ of the interface along any sequence $\varepsilon_n \downarrow 0$.

The rest of this section is organized as follows. In Subsection 3.2 we derive a uniform upper bound on the expected number of boundaries $I_1(X^\varepsilon_\infty)$, and in Subsection 3.3, we use this to prove weak continuity of the map $\varepsilon \mapsto \nu_\varepsilon$ with respect to the product topology on $\hat{S}^{01}_{\rm int}$. In Subsection 3.4, we establish the equilibrium equation $\mathbb{E}[G^0 h^0(X^0_\infty)] = 0$, which is rather nontrivial in itself since the function h^0 is not bounded. Once this is done, however, the proof of Theorem 1.3 in Subsection 3.5 is quite short.

3.2 Bound on the number of boundaries

The following lemma is our basic tool for controlling the number of boundaries of an equilibrium interface. We note that in Proposition 3.7 below, when we establish the equilibrium equation $\mathbb{E}[G^0 h^0(X^0_\infty)] = 0$, we will see that for $\varepsilon = 0$, (3.2) is in fact an equality and this fact will be key to our proof of Theorem 1.3.

Lemma 3.1 (Bound on k-boundaries) *For $\varepsilon \in [0,1)$, let X^ε_∞ denote a random variable with law ν_ε as in Subsection 3.1. Then*

$$\mathbb{E}\big[\sum_{k=1}^{\infty} a_{\rm s}(k) I_k(X^\varepsilon_\infty)\big] \le \tfrac{1}{2}\sigma^2, \tag{3.2}$$

where I_k is defined in (2.4), $a_{\rm s}(k) = \frac{1}{2}\big(a(k) + a(-k)\big)$ and $\sigma^2 = \sum_k a(k)k^2$.

Proof. Let the biased voter model $(X^\varepsilon_t)_{t\ge 0}$ start from the heaviside initial state $X^\varepsilon_0 = x_0$ as in (1.4). For $\varepsilon \in (0,1)$, recall from (2.3) that

$$h^\varepsilon(x) = \sum_{i>j} (1-\varepsilon)^{\sum_{n<j} x(n)} 1_{\{x(j,i)=10\}}. \tag{3.3}$$

Hence $h^\varepsilon(X^\varepsilon_0) \equiv 0$. By Lemma 2.2 and Lemma 2.5,

$$\int_0^t \mathbb{E}\big[\sum_{k=1}^{\infty} a_{\rm s}(k)\big(k^2 - I_k(X^\varepsilon_s)\big)\big]{\rm d}s \ge \int_0^t \mathbb{E}\big[G^\varepsilon h^\varepsilon(X^\varepsilon_s)\big]{\rm d}s \ge 0 \qquad (t>0). \tag{3.4}$$

Dividing both sides of (3.4) by t and then letting $t \to \infty$, we arrive at (3.2) by Fatou's lemma, since $\overline{X}^\varepsilon_t$ converges weakly to $\overline{X}^\varepsilon_\infty$. ∎

Lemma 3.2 (Bound on 1-boundaries) *There exists a constant $C < \infty$ such that*

$$\sup_{\varepsilon\in[0,1)} \mathbb{E}\big[I_1(X^\varepsilon_\infty)\big] \le C. \tag{3.5}$$

Proof. Fix $t > 0$, and choose $k \ge 1$ such that $a_{\rm s}(k) > 0$. For a biased voter model started in an initial state x such that $x(i,i+1) = 10$, using the irreducibility of the kernel a, it is easy to see that there is a positive probability p that the 1 at i spreads its type to $i-k+1$ while leaving the zero at $i+1$ as it is. Since this event only requires 1's to spread, this probability is uniform in the bias ε. Therefore, letting $(X^\varepsilon_t)_{t\ge 0}$ denote the biased voter model started in the initial state ν_ε,

$$\mathbb{E}\big[I^{10}_1(X^\varepsilon_0)\big] = \sum_i \mathbb{E}\big[1_{X^\varepsilon_0(i,i+1)=10}\big] \le \sum_i \frac{1}{p}\mathbb{E}\big[1_{X^\varepsilon_t(i-k+1,i+1)=10}\big] = \frac{1}{p}\mathbb{E}\big[I^{10}_k(X^\varepsilon_t)\big]. \tag{3.6}$$

Since the law of X^ε_t modulo translations does not depend on t, the claim now follows from Lemma 3.1. ∎

3.3 Continuity with respect to the product topology

In this subsection we prove the following theorem.

Theorem 3.3 (Continuity with respect to the product topology) *Assume that the kernel a is irreducible and satisfies $\sum_k a(k)k^2 < \infty$. Equip $\{0,1\}^{\mathbb{Z}}$ with the product topology. Then the map $[0,1) \ni \varepsilon \mapsto \nu_\varepsilon$ is continuous with respect to weak convergence of probability measures on $\{0,1\}^{\mathbb{Z}}$.*

To prepare for the proof of Theorem 3.3, we need a few lemmas. Let $(X^\varepsilon_t)_{t\geq 0}$ denote the biased voter model started in the initial state ν_ε. At time $t \geq 0$, denote the position of the leftmost 1 and rightmost zero by l^ε_t and r^ε_t, that is,

$$l^\varepsilon_t := \min\{i : X^\varepsilon_t(i) = 1\} \quad \text{and} \quad r^\varepsilon_t := \max\{i : X^\varepsilon_t(i) = 0\} \qquad (t \geq 0). \tag{3.7}$$

For $i \in \mathbb{Z}$, define a shift operator θ_i

$$\theta_i(x)(j) := x(i + j) \qquad (i, j \in \mathbb{Z},\ x \in S^{01}_{\rm int}). \tag{3.8}$$

Then $\theta_{l_t}(X^\varepsilon_t)$ has law ν_ε for all $t \geq 0$.

Our strategy for proving Theorem 3.3 is as follows. Fix $\varepsilon_n, \varepsilon^* \in [0, 1)$ such that $\varepsilon_n \to \varepsilon^*$. Since the space $\{0, 1\}^{\mathbb{Z}}$ is compact, tightness comes for free so by going to a subsequence if necessary, we can assume that $\nu_{\varepsilon_n} \Rightarrow \nu^*$ for some probability measure ν^* on $\{0, 1\}^{\mathbb{Z}}$. Using Lemma 3.2, we see that each subsequential limit ν^* is concentrated on $S^{01}_{\rm int}$. Using convergence of the generators, general arguments tell us that $(X^{\varepsilon_n}_t)_{t\geq 0}$ converges to the voter model $(X^*_t)_{t\geq 0}$ with initial law ν^*. This is Lemma 3.4 below. Next, in Lemma 3.6, we show that for any fixed t, the family $(l^\varepsilon_t)_{\varepsilon\in[0,1)}$ is tight. Using this, we can show that ν_* is an invariant law for the process modulo translations. Since this invariant law is unique, $\nu_* = \nu_{\varepsilon^*}$, establishing the continuity of the map $\varepsilon \mapsto \nu_\varepsilon$.

Lemma 3.4 (Convergence as a process) *Equip $\{0, 1\}^{\mathbb{Z}}$ with the product topology. Assume that $\varepsilon_n \to \varepsilon^* \in [0, 1)$ are such that $\nu_{\varepsilon_n} \Rightarrow \nu^*$ for some probability measure ν^* on $\{0, 1\}^{\mathbb{Z}}$. Then*

$$\mathbb{P}\big[X^{\varepsilon_n}_t \in \cdot\,\big] \underset{n\to\infty}{\Longrightarrow} \mathbb{P}\big[X^*_t \in \cdot\,\big] \qquad (t \geq 0), \tag{3.9}$$

where $X^{\varepsilon_n}_0$ has law ν^{ε_n}, and $(X^*_t)_{t\geq 0}$ is the voter model with bias ε^* and initial law ν^*.

Proof. By [Lig85, Theorem 3.9], for each $\varepsilon \in [0, 1)$, the generator G^ε in (1.1) is well-defined for any function $f \in \mathcal{D}$ with

$$\mathcal{D} := \Big\{f : \sum_{i\in\mathbb{Z}} \sup_{x\in\{0,1\}^{\mathbb{Z}}} \big|f(x + e_i) - f(x)\big| < \infty\Big\}, \tag{3.10}$$

and the closure of the generator G^ε with domain $\mathcal{D}$ generates a Feller semigroup. By [Kal97, Theorem 17.25], for (3.9) it suffices to check that $\|G^{\varepsilon_n} f - G^{\varepsilon^*} f\|_\infty \to 0$ for all $f \in \mathcal{D}$, where $\|\cdot\|_\infty$ denotes the supremum norm. This follows by writing

$$\begin{aligned}\big|G^{\varepsilon_n} f(x) - G^{\varepsilon^*} f(x)\big| &= |\varepsilon_n - \varepsilon^*| \cdot \Big|\sum_{k\neq 0} a(k) \sum_i 1_{x(i-k,i)=01}\big\{f(x - e_i) - f(x)\big\}\Big| \\ &\leq |\varepsilon_n - \varepsilon^*| \cdot \sum_{k\neq 0} a(k) \sum_{i\in\mathbb{Z}} \big|f(x - e_i) - f(x)\big|.\end{aligned} \tag{3.11}$$

∎

Our next aim is to show that the position l^ε_t of the leftmost 1 is tight in the bias parameter ε. We need the following simple fact.

Lemma 3.5 (Stationary increments) *Let X and Y be real random variables that are equal in distribution, and assume that $\mathbb{E}\big[(Y-X)\vee 0\big] < \infty$. Then $\mathbb{E}\big[|Y-X|\big] < \infty$ and $\mathbb{E}[Y-X] = 0$.*

Proof. It suffices to show that $\mathbb{E}\big[(X-Y)\vee 0\big] = \mathbb{E}\big[(Y-X)\vee 0\big]$. For any real random variable Z and constant $c > 0$, write $Z^c := Z \wedge c$ and $Z_c := Z \vee (-c)$. Then $\mathbb{E}[X^n_n - Y^n_n] = 0$ and hence $\mathbb{E}\big[(X^n_n - Y^n_n) \vee 0\big] = \mathbb{E}\big[(Y^n_n - X^n_n) \vee 0\big]$. By monotone convergence

$$\mathbb{E}\big[(X^n_n - Y^n_n) \vee 0\big] = \mathbb{E}\big[1_{\{-n<X\}} 1_{\{Y<n\}} (X^n - Y_n) \vee 0\big] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\big[(X - Y) \vee 0\big], \tag{3.12}$$

and similarly $\mathbb{E}\big[(Y^n_n - X^n_n) \vee 0\big]$ converges to $\mathbb{E}\big[(Y - X) \vee 0\big]$. ∎

Lemma 3.6 (Expected displacement of the leftmost one) *There exists a constant $C < \infty$ such that uniformly in $\varepsilon \in [0,1)$ and $t \geq 0$,*

$$\mathbb{E}\big[|l_t^\varepsilon|\big] \leq Ct. \tag{3.13}$$

Proof. We first lower bound l_t^ε. Since for any $\varepsilon \in [0,1)$, the rate that $(l_t^\varepsilon)_{t\geq 0}$ jumps to the left by k ($k > 0$) is given by $\sum_{n\geq 0} 1_{\{X_t^\varepsilon(l_t^\varepsilon+n)=1\}} a(-n-k)$, we can couple it with a uni-directional random walk $(S_t)_{t\geq 0}$ started in $S_0 = 0$ with increment rate $q(k) := \sum_{n\geq 0} a(-n-k)$ for $k < 0$ and $q(k) = 0$ for $k > 0$, such that $S_t \leq l_t^\varepsilon$ for all $t \geq 0$ almost surely. This gives the estimate

$$\mathbb{E}\big[(-l_t^\varepsilon) \vee 0\big] \leq \mathbb{E}[|S_t|] = t\sum_{k\geq 0} k \sum_{n\geq 0} a(-n-k) = t\sum_{k\geq 1} a(-k)\tfrac{1}{2}k(k+1). \tag{3.14}$$

The same argument applied to the rightmost zero gives

$$\mathbb{E}\big[(r_t^\varepsilon - r_0^\varepsilon) \vee 0\big] \leq t\sum_{k\geq 1} a(k)\tfrac{1}{2}k(k+1). \tag{3.15}$$

Together, these estimates show that

$$\mathbb{E}\big[\big\{(r_t^\varepsilon - l_t^\varepsilon) - r_0^\varepsilon\big\} \vee 0\big] < \infty, \tag{3.16}$$

so we can apply Lemma 3.5 to the equally distributed random variables $(r_t^\varepsilon - l_t^\varepsilon)$ and r_0^ε to conclude that

$$\mathbb{E}\big[(r_t^\varepsilon - l_t^\varepsilon) - r_0^\varepsilon\big] = 0. \tag{3.17}$$

Formulas (3.14) and (3.15) show that $\mathbb{E}[l_t^\varepsilon]$ is well-defined (but may be $+\infty$) and also $\mathbb{E}[r_t^\varepsilon - r_0^\varepsilon]$ is well-defined (but may be $-\infty$), so (3.17) tells us that $\mathbb{E}[l_t^\varepsilon] = \mathbb{E}\big[r_t^\varepsilon - r_0^\varepsilon\big]$. This gives

$$\begin{aligned}\mathbb{E}\big[l_t^\varepsilon \vee 0\big] - \mathbb{E}\big[(-l_t^\varepsilon) \vee 0\big] &= \mathbb{E}[l_t^\varepsilon] = \mathbb{E}[r_t^\varepsilon - r_0^\varepsilon] \\ &= \mathbb{E}\big[(r_t^\varepsilon - r_0^\varepsilon) \vee 0\big] - \mathbb{E}\big[(r_0^\varepsilon - r_t^\varepsilon) \vee 0\big] \leq \mathbb{E}\big[(r_t^\varepsilon - r_0^\varepsilon) \vee 0\big],\end{aligned} \tag{3.18}$$

and hence

$$\mathbb{E}\big[|l_t^\varepsilon|\big] = \mathbb{E}\big[l_t^\varepsilon \vee 0\big] + \mathbb{E}\big[(-l_t^\varepsilon) \vee 0\big] \leq \mathbb{E}\big[(r_t^\varepsilon - r_0^\varepsilon) \vee 0\big] + 2\mathbb{E}\big[(-l_t^\varepsilon) \vee 0\big] \leq Ct \tag{3.19}$$

with $C = \sum_{k\geq 1} \big[a(-k) + \frac{1}{2}a(k)\big]k(k+1)$. ∎

Proof of Theorem 3.3. Let $\varepsilon_n, \varepsilon^* \in [0,1)$ and $\varepsilon_n \to \varepsilon^*$. We need to show that $\nu_{\varepsilon_n} \Rightarrow \nu_{\varepsilon^*}$. Since the space $\{0,1\}^{\mathbb{Z}}$ is compact under the product topology, tightness comes for free, so without loss of generality we may assume that $\nu_{\varepsilon_n} \Rightarrow \nu^*$ for some probability measure ν^* on $\{0,1\}^{\mathbb{Z}}$. By Lemma 3.2, ν^* is concentrated on $S^{01}_{\rm int}$. Let $(X_t^{\varepsilon_n})_{t\geq 0}$ denote the voter model with bias ε_n started in ν_{ε_n}. Using again the compactness of $\{0,1\}^{\mathbb{Z}}$, as well as the fact that by Lemma 3.6, the laws of the random variables $(l_t^{\varepsilon_n})^{n\geq 1}$ are tight, going to a subsequence if necessary, we can without loss of generality assume that $(X_t^{\varepsilon_n}, l_t^{\varepsilon_n})$ converges in law to some random variable (X_t^*, l_t^*). By Skorohod's representation theorem, we can find a coupling such that the convergence is almost sure. From this, it is easy to see that $l_t^* = \min\{i : X_t^*(i) = 1\}$. By Lemma 3.4, X_t^* is distributed as the voter model with bias ε^*, started in ν^*, and evaluated at time t. Since

$$\theta_{l_t^{\varepsilon_n}}\big(X_t^{\varepsilon_n}\big) \underset{n\to\infty}{\longrightarrow} \theta_{l_t^*}\big(X_t^*\big) \quad \text{a.s.} \tag{3.20}$$

and the left-hand side of this equation has law ν_{ε_n} for each $t \geq 0$, we see that ν^* is an invariant law for the voter model with bias ε^* as seen from the leftmost one. By the uniqueness of the latter (here we are using 2.1), we conclude that $\nu^* = \nu_{\varepsilon^*}$. ∎

3.4 Equilibrium equation

In this subsection, we will only consider the unbiased voter model, and the main purpose is to establish the equilibrium equation (3.22) below, which shows that (3.2) holds with equality if $\varepsilon = 0$. For brevity, throughout this section we drop the superscript indicating the bias, for example, the generator G^0 is abbreviated by G. Let X_∞ be a random variable taking values in the space $\hat{S}^{01}_{\rm int}$ from (3.1) with law ν_0, i.e., $\min\{i : X_\infty(i) = 1\} = 0$ and the law of X_∞ is stationary for the unbiased voter model modulo translations, that is, the law of $\overline{X}_\infty$ is $\overline{\nu}_0$.

Proposition 3.7 (Equilibrium equation) *Assume that the kernel a has finite second moment $\sigma^2 = \sum_k a(k)k^2 < \infty$. Then, the following equilibrium equation for the voter model holds,*

$$\mathbb{E}\big[Gh(X_\infty)\big] = 0, \tag{3.21}$$

where $h = h^0$ is the number of inversions (2.2)*. Or equivalently, by the expression of Gh in* (2.5)*, we have*

$$\mathbb{E}\Big[\sum_{k=1}^{\infty} a_{\rm s}(k) I_k(X_\infty)\Big] = \tfrac{1}{2}\sigma^2, \tag{3.22}$$

where $a_{\rm s}(k) = \frac{1}{2}\big(a(k) + a(-k)\big)$ and $I_k(x) = |\{i : x(i) \neq x(i+k)\}|$.

We briefly explain our proof strategy. If $h(X_t) - \int_0^t Gh(X_s){\rm d}s$ were a martingale for any deterministic initial configuration X_0, and if $\mathbb{E}[h(X_\infty)]$ and $\mathbb{E}[|Gh(X_\infty)|]$ were finite, then the equilibrium equation (3.21) would follow. However, two difficulties arise in this approach. As shown in Remark 2.10, we can only prove that $h(X_t) - \int_0^t Gh(X_s){\rm d}s$ is a martingale when $a(\cdot)$ has finite third moment, as otherwise there is no control on the expected number of inversions $\mathbb{E}[h(X_t)]$. Worse still, since h is bounded from below by the length L of the interface and $\mathbb{E}[L(X_\infty)] = \infty$ by [CD95, Theorem 6] or [BMSV06, Theorem 1.4], we have $\mathbb{E}[h(X_\infty)] = \infty$. To bypass these difficulties, we will show that the equilibrium equation (3.21) holds for h_M in place of h, where

$$h_M(x) = \big|\{(i,j) : 0 < j - i \leq M,\ x(i,j) = 10\}\big| \tag{3.23}$$

as defined in (2.71) is a truncation of $h(x)$. We will then take the limit as $M \to \infty$ to deduce (3.21). To deduce this last convergence, we will use the fact that the expected number of 1-boundaries $\mathbb{E}[I_1(X_\infty)]$ is finite, by Lemma 3.2.

Our first step is to do a generator calculation for h_M. Recall formula (2.8) for Gh, where h is the number of inversions. The next lemma identifies Gh_M.

Lemma 3.8 *For any $x \in S^{01}_{\rm int}$ and $M \in \mathbb{N}$, we have*

$$Gh_M(x) = \sum_{k=1}^{\infty} a_{\rm s}(k)\big(k^2 - I_k(x)\big) + \sum_{n=1}^{\infty} A(n) I^{10}_{M+n}(x) - \sum_{n=1}^{\infty} A(n) I^{10}_{M-1-n}(x), \tag{3.24}$$

where $A(n) := \sum_{k=n}^{\infty} \big(a(k) + a(-k)\big) = 2\sum_{k=n}^{\infty} a_{\rm s}(k)$.

Proof. We use the generator decomposition $G = \sum_{k\neq 0} a(k)G_k$ in (2.18), and separately calculate $G_k h_M$ for $k > 0$ and $k < 0$.

For $k > 0$, to calculate $G_k h_M(x)$, we consider all triples $(i, j, j-k)$ with $|i-j| \leq M$, where an inversion in $x(i,j)$ is either created or destroyed because $x(j)$ changes its value to that of

$x(j-k)$. Therefore,

$$\begin{aligned} G_k h_M(x) =& \sum_i 1_{\{x(i)=1\}}\Big\{ -\sum_{j=i+1}^{i+M} 1_{\{x(j-k,j)=10\}} + \sum_{j=i+1}^{i+M} 1_{\{x(j-k,j)=01\}}\Big\} \\ &+ \sum_i 1_{\{x(i)=0\}}\Big\{ \sum_{j=i-M}^{i-1} 1_{\{x(j-k,j)=10\}} - \sum_{j=i-M}^{i-1} 1_{\{x(j-k,j)=01\}}\Big\} \\ =& \sum_i 1_{\{x(i)=1\}} \sum_{j=i+1}^{i+M} \big(1_{\{x(j-k)=0\}} - 1_{\{x(j)=0\}}\big) \\ &+ \sum_i 1_{\{x(i)=0\}} \sum_{j=i-M}^{i-1} \big(1_{\{x(j-k)=1\}} - 1_{\{x(j)=1\}}\big), \end{aligned} \tag{3.25}$$

where in the last equality we used (2.28). To further simplify this, we observe that for any $a, b_1, b_2, c \in \mathbb{Z}$ with $a < b_1, b_2 < c$,

$$1_{\{a<j\le b_1\}} - 1_{\{b_2<j\le c\}} = 1_{\{a<j\le b_2\}} - 1_{\{b_1<j\le c\}} \qquad (j \in \mathbb{Z}). \tag{3.26}$$

Applying this to $a = i-k$, $b_1 = i$, $b_2 = i+M-k$, and $c = i+M$, we get

$$\begin{aligned} &\sum_{j=i+1}^{i+M} \big(1_{\{x(j-k)=0\}} - 1_{\{x(j)=0\}}\big) = \sum_{j\in\mathbb{Z}} 1_{\{x(j)=0\}}\big(1_{\{i-k<j\le i+M-k\}} - 1_{\{i<j\le i+M\}}\big) \\ &\quad = \sum_{j\in\mathbb{Z}} 1_{\{x(j)=0\}}\big(1_{\{i-k<j\le i\}} - 1_{\{i+M-k<j\le i+M\}}\big) = \sum_{n=0}^{k-1} \big(1_{\{x(i-n)=0\}} - 1_{\{x(i+M-n)=0\}}\big), \end{aligned} \tag{3.27}$$

and similarly,

$$\begin{aligned} &\sum_{j=i-M}^{i-1} \big(1_{\{x(j-k)=1\}} - 1_{\{x(j)=1\}}\big) = \sum_{j\in\mathbb{Z}} 1_{\{x(j)=1\}}\big(1_{\{i-k-M\le j<i-k\}} - 1_{\{i-M\le j<i\}}\big) \\ &\quad = \sum_{j\in\mathbb{Z}} 1_{\{x(j)=1\}}\big(1_{\{i-k-M\le j<i-M\}} - 1_{\{i-k\le j<i\}}\big) = \sum_{n=1}^{k} \big(1_{\{x(i-M-n)=1\}} - 1_{\{x(i-n)=1\}}\big). \end{aligned} \tag{3.28}$$

Substituting this into the right-hand side of (3.25), and then using the notation I^{01}_k, I^{10}_k and I_k as in (2.4), we can rewrite (3.25) as

$$\begin{aligned} G_k h_M(x) =& \sum_i \sum_{n=0}^{k-1} \big(1_{\{x(i-n,i)=01\}} - 1_{\{x(i,i+M-n)=10\}}\big) \\ &+ \sum_i \sum_{n=1}^{k} \big(1_{\{x(i-M-n,i)=10\}} - 1_{\{x(i-n,i)=10\}}\big) \\ =& \sum_{n=1}^{k-1} I^{01}_n(x) - \sum_{n=0}^{k-1} I^{10}_{M-n}(x) - \sum_{n=1}^{k} I^{10}_n(x) + \sum_{n=1}^{k} I^{10}_{M+n}(x) \\ =& \tfrac{1}{2}\big(k^2 - I_k(x)\big) + \sum_{n=1}^{k} \big(I^{10}_{M+n}(x) - I^{10}_{M-1-n}(x)\big). \end{aligned} \tag{3.29}$$

where in the last equality we applied (2.37) to $\sum_{n=1}^{k-1}\big(I^{01}_n(x) - I^{10}_n(x)\big) - I^{10}_k(x)$.

Using symmetry, we now also easily obtain a formula for $G_k h_M$ when $k < 0$. For any $x \in S^{01}_{\rm int}$, define $x' \in S^{01}_{\rm int}$ by $x'(i) := 1 - x(-i)$ $(i \in \mathbb{Z})$. Then, for any function $f : S^{01}_{\rm int} \to \mathbb{R}$, one has $G_k f(x) = G_{-k} f'(x')$, where $f'(x) := f(x')$ $(x \in S^{01}_{\rm int})$. We observe that $I_k(x) = I_k(x')$

and hence also $h_M(x) = \sum_{k=1}^{M} I_k(x)$ is symmetric in the sense that $h_M(x) = h_M(x')$ ($x \in S_{\text{int}}^{01}$). Combining these observations with (3.29), we obtain that for any $k \neq 0$

$$G_k h_M(x) = \tfrac{1}{2}\big(k^2 - I_{|k|}(x)\big) + \sum_{n=1}^{|k|} \big(I_{M+n}^{10}(x) - I_{M-1-n}^{10}(x)\big). \tag{3.30}$$

Inserting this into $G = \sum_{k\neq 0} a(k) G_k$, we have

$$Gh_M(x) = \sum_{k=1}^{\infty} a_{\text{s}}(k)\big(k^2 - I_k(x)\big) + 2\sum_{k=1}^{\infty} a_{\text{s}}(k) \sum_{n=1}^{k} \big(I_{M+n}^{10}(x) - I_{M-1-n}^{10}(x)\big). \tag{3.31}$$

Interchanging the summation order, we obtain (3.24). ∎

We are now ready to prove Proposition 3.7.

Proof of Proposition 3.7. Let $(X_t)_{t\geq 0}$ be the voter model starting from some configuration $x \in S_{\text{int}}^{01}$. Under the second moment assumption, by (2.76) for any $t > 0$,

$$\mathbb{E}\big[h_M(X_t)\big] - h_M(x) = \int_0^t \mathbb{E}\big[Gh_M(X_s)\big]\mathrm{d}s. \tag{3.32}$$

Assume for the moment that both h_M and Gh_M are absolutely integrable with respect to the law of X_∞. Then we can integrate both sides of (3.32) with respect to the invariant law to get

$$\mathbb{E}\big[Gh_M(X_\infty)\big] = 0. \tag{3.33}$$

By letting $M \to \infty$, we will see in the following that (3.33) implies (3.21), the equilibrium equation for the voter model. Recalling the expression of $Gh_M(x)$ in (3.24) and $Gh(x)$ in (2.8), we obtain from (3.33) that

$$\begin{aligned}\mathbb{E}\big[Gh(X_\infty)\big] &= \mathbb{E}\Big[\sum_{k=1}^{\infty} a_{\text{s}}(k)\big(k^2 - I_k(X_\infty)\big)\Big] \\ &= \mathbb{E}\Big[\sum_{n=1}^{\infty} A(n) I_{M-1-n}^{10}(X_\infty) - \sum_{n=1}^{\infty} A(n) I_{M+n}^{10}(X_\infty)\Big],\end{aligned} \tag{3.34}$$

where $A(n) = 2\sum_{k=n}^{\infty} a_{\text{s}}(k)$. For $n \geq M$, by (2.4) and (2.37), we have

$$I_{-(n+1-M)}^{10}(x) = I_{n+1-M}^{01}(x) = I_{n+1-M}^{10}(x) + (n+1-M). \tag{3.35}$$

Therefore, by Lemma 2.7,, we can bound the difference in the expectation in (3.34) by

$$\begin{aligned}&\Big|\sum_{n=1}^{\infty} A(n) I_{M-1-n}^{10}(X_\infty) - \sum_{n=1}^{\infty} A(n) I_{M+n}^{10}(X_\infty)\Big| \\ &\quad\leq \sum_{n=1}^{M-1} A(n)\big|I_{M-1-n}^{10}(X_\infty) - I_{M+n}^{10}(X_\infty)\big| \\ &\qquad + \sum_{n=M}^{\infty} A(n)\big|I_{n+1-M}^{10}(X_\infty) + (n+1-M) - I_{M+n}^{10}(X_\infty)\big| \\ &\quad\leq \sum_{n=1}^{M-1} A(n)(2n+1) I_1^{10}(X_\infty) + \sum_{n=M}^{\infty} A(n)\big\{n+1-M+(2M-1)I_1^{10}(X_\infty)\big\} \\ &\quad\leq \sum_{n=1}^{\infty} 3nA(n)\big(1 + I_1^{10}(X_\infty)\big).\end{aligned} \tag{3.36}$$

Due to the second moment assumption, $\sum_{n=1}^{\infty} nA(n)$ is finite, so applying Lemma 3.2 we see that the right-hand side of (3.36) is bounded in expectation. As a result, the term in the expectation on the right-hand side of (3.34) is in absolute value bounded by an integrable random variable, uniformly in M. If this term moreover converges to zero pointwise as M tends to ∞, then applying the dominated convergence theorem to (3.34), we will obtain the equilibrium equation

$$\mathbb{E}\big[Gh(X_\infty)\big] = 0. \tag{3.37}$$

To see the pointwise convergence, note that for every $x \in S^{01}_{\rm int}$, there exists some M_x such that $I_k(x) = 0$ for all $|k| > M_x$, and thus when $M > M_x + 1$,

$$\sum_{n=1}^{\infty} A(n) I^{10}_{M-1-n}(x) - \sum_{n=1}^{\infty} A(n) I^{10}_{M+n}(x) = \sum_{k=-M_x}^{M_x} A(M-1-k) I_k(x), \tag{3.38}$$

where the right-hand side decreases to 0 since $A(M-1-k) \downarrow 0$ as M tends to ∞.

To complete the proof, it remains to show, for fixed M, the absolute integrability of h_M and Gh_M with respect to the invariant law. For the nonnegative function h_M, by Lemmas 2.7 and 3.2,

$$\mathbb{E}\big[h_M(X_\infty)\big] = \mathbb{E}\Big[\sum_{k=1}^{M} I^{10}_k(X_\infty)\Big] \leq \mathbb{E}\Big[\sum_{k=1}^{M} k I^{10}_1(X_\infty)\Big] < \infty. \tag{3.39}$$

By using the expression of Gh_M in (3.24), Lemma 2.7, the fact that $\sum_{n=1}^{\infty} A(n)n < \infty$ since $\sum_k a(k)k^2 < \infty$, and Lemma 3.2, it is also not hard to see that $\mathbb{E}\big[|Gh_M(X_\infty)|\big] < \infty$. ∎

3.5 Proof of Theorem 1.3

For each $\varepsilon \geq 0$, let X^ε_∞ denote a random variable taking values in the space $\hat{S}^{01}_{\rm int}$ from (3.1) with law ν_ε. Since there is a one-to-one correspondence between $\hat{S}^{01}_{\rm int}$ and $\overline{S}^{01}_{\rm int}$, it suffices to show that as $\varepsilon \downarrow 0$, the measures ν_ε converge weakly to ν_0 with respect to the discrete topology on $\hat{S}^{01}_{\rm int}$. By Theorem 3.3, the measures ν_ε converge weakly to ν_0 with respect to the product topology on $\{0,1\}^{\mathbb{Z}}$. To improve this to convergence with respect to the discrete topology, it suffices to show that for any $\varepsilon_n \downarrow 0$, the laws of the random variables

$$\big(r^{\varepsilon_n}_\infty\big)_{n\geq 1} \tag{3.40}$$

are tight, where as in (3.7), we let $r^\varepsilon_\infty := \max\{i : X^\varepsilon_\infty(i) = 0\}$ denote the position of the rightmost zero of X^ε_∞.

Assume that for some $\varepsilon_n \downarrow 0$, the laws of the random variables in (3.40) are not tight. Then going to a subsequence if necessary, we can find $\delta > 0$ and $(m_N)_{N\geq 1}$ such that

$$\mathbb{P}\big[r^{\varepsilon_n}_\infty > N\big] > \delta \qquad (n \geq m_N). \tag{3.41}$$

For $x \in S^{01}_{\rm int}$ and $N \in \mathbb{Z}$, let

$$I^N_k(x) := \big|\{i \leq N : x(i) \neq x(i+k)\}\big|. \tag{3.42}$$

Since $X^\varepsilon_\infty(r^\varepsilon_\infty) \neq X^\varepsilon_\infty(r^\varepsilon_\infty + k)$ for all $k \geq 1$, by Lemma 3.1 and (3.41), we see that

$$\mathbb{E}\big[\sum_{k=1}^{\infty} a_{\rm s}(k) I^N_k(X^{\varepsilon_n}_\infty)\big] \leq \tfrac{1}{2}\sigma^2 - A\delta \qquad (n \geq m_N), \tag{3.43}$$

where $A = \sum_{k=1}^{\infty} a_{\rm s}(k) > 0$. Letting $n \to \infty$, using weak convergence with respect to the product topology (Theorem 3.3), we find that

$$\mathbb{E}\big[\sum_{k=1}^{\infty} a_{\rm s}(k) I_k^N(X_\infty^0)\big] \leq \tfrac{1}{2}\sigma^2 - A\delta. \tag{3.44}$$

Since $I_k^N \uparrow I_k$ as $N \to \infty$, this contradicts the equilibrium equation (3.22) from Proposition 3.7. ∎

Acknowledgement J.M. Swart is sponsored by grant 16-15238S of the Czech Science Foundation (GA CR). R. Sun and J. Yu are partially supported by NUS grant R-146-000-220-112. We would like to thank the Institute for Mathematical Sciences at NUS for hospitality during the program *Genealogies of Interacting Particle Systems*, where part of this work was done.

References

[AMPV10] E. Andjel, T. Mountford, L.P.R. Pimentel, and D. Valesin. Tightness for the interface of the one-dimensional contact process. *Bernoulli* 16(4), 909–925, 2010.

[BMSV06] S. Belhaouari, T. Mountford, R. Sun, and G. Valle. Convergence results and sharp estimates for the voter model interfaces. *Electron. J. Probab.* 11, 768–801, 2006.

[BMV07] S. Belhaouari, T. Mountford, and G. Valle. Tightness for the interfaces of one-dimensional voter models. *Proc. London Math. Soc.* 94(3), 421–442, 2007.

[CD95] J.T. Cox and R. Durrett. Hybrid zones and voter model interfaces. *Bernoulli* 1(4), 343–370, 1995.

[FINR04] L.R.G. Fontes, M. Isopi, C.M. Newman, and K. Ravishankar. The Brownian web: characterization and convergence. *Ann. Probab.* 32(4), 2857–2883, 2004.

[Kal97] O. Kallenberg. *Foundations of Modern Probability.* Springer, New York, 1997.

[Lig85] T. M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[MPW17] M. Menshikov, S. Popov, A. Wade. *Non-homogeneous Random Walks: Lyapunov Function Methods for Near-critical Stochastic Systems.* Cambridge University Press, 2017.

[MV16] M. Mountford and D. Valesin. Functional central limit theorem for the interface of the symmetric multitype contact process. *ALEA* 13(1), 481–519, 2016.

[NRS05] C.M. Newman, K. Ravishankar, R. Sun. Convergence of coalescing nonsimple random walks to the Brownian web. *Electron. J. Prob.* 10, 21–60, 2005.

[SS08a] A. Sturm and J.M. Swart. Voter models with heterozygosity selection. *Ann. Appl. Probab.* 18(1), 59–99, 2008.

[SS08b] A. Sturm and J.M. Swart. Tightness of voter model interfaces. *Electron. Commun. Probab.* 13, No. 16, 165–174, 2008.

[SS08c] R. Sun and J.M. Swart. The Brownian net. *Ann. Probab.* 36, 1153–1208, 2008.

[SS15] A. Sturm and J.M. Swart. A particle system with cooperative branching and coalescence. *Ann. Appl. Probab.* 25(3), 1616–1649, 2015.

[SSS17] E. Schertzer, R. Sun, and J.M. Swart. The Brownian web, the Brownian net, and their universality. *Advances in Disordered Systems, Random Processes and Some Applications*, 270–368, Cambridge University Press, 2017.

[Sch77] D. Schwartz. Applications of duality to a class of Markov processes. *Ann. Probab.* 5, 522–532, 1977.

[Val10] D. Valesin. Multitype contact process on $\mathbb{Z}$: extinction and interface. *Electron. J. Probab.* 15, No. 73, 2220–2260, 2010.

[WB72] T. Williams and R. Bjerknes. Stochastic model for abnormal clone spread through epithelial basal layer. *Nature* 236, 19–21, 1972.