Electronic Journal of Probability

Electron. J. Probab. ( ), article no. , 1–80.
ISSN: 1083-6489 https://doi.org/

A complete convergence theorem for the q-voter model and other voter model perturbations in two dimensions

Ted Cox* Ed Perkins†

Abstract

The q-voter model is a spin-flip system in which the rate of flipping to type i is given by the qth power of the proportion of nearest neighbours in type i for $i = 0, 1$. If $q = 1$ it reduces to the classical voter model. We show that in the critical 2-dimensional case, for $q < 1$ and close enough to 1, for any initial state as $t \to \infty$ the system converges weakly to a mixture of all 0's, all 1's, and a unique invariant law which contains infinitely many sites of both types. This follows as a special case of a general theorem which proves a similar "complete convergence theorem" for cancellative, monotone, finite range voter model perturbations on $\mathbb{Z}^2$ providing a certain parameter, Θ_3, is strictly positive. Similar results follow for the affine and geometric voter models and Lotka-Volterra models, all for parameter values close to that giving the voter model. This kind of asymptotic behavior is quite different from that of the 2-dimensional voter model itself, which undergoes clustering, and converges to a mixture of all 0's and all 1's.

The above parameter Θ_3 has an explicit expression in terms of asymptotic coalescing probabilities of 2-dimensional random walk, and we give a rather simple sufficient condition for it to be strictly positive. An important step in the proof is to establish weak convergence of the rescaled spin-flip systems to super-Brownian motion with drift Θ_3. In fact, a convergence result is proved under weaker hypotheses which includes all known such results for 2-dimensional voter model perturbations and a number of new ones, including a rescaled limit theorem for the q-voter model where $q \uparrow 1$ with the rescaling.

Keywords: Complete convergence theorem; q-voter model; interacting particle system; voter model perturbation; cancellative processes; super-Brownian motion; two-dimensional random walk.
MSC2020 subject classifications: Primary 60K35; 82C22, Secondary 60F99; 91D30.
Submitted to EJP on , final version accepted on .

*Syracuse University, United States of America. E-mail: jtcox@syr.edu
†The University of British Columbia, Canada. E-mail: perkins@math.ubc.ca

1 Introduction and main results

1.1 The q-voter model

To define the model, let $q \geq 0$ and let $\mathcal{N} \subset \mathbb{Z}^d$ be a non-empty finite symmetric (about the origin) set not containing 0 such that the uniform distribution on $\mathcal{N}$ is an irreducible kernel (the group generated by $\mathcal{N}$ is $\mathbb{Z}^d$), and for some $\sigma^2 > 0$,

$$\sum_{z \in \mathcal{N}} z_i z_j / |\mathcal{N}| = \delta_{ij}\sigma^2 \text{ for all } i, j \leq d. \tag{1.1}$$

Here $|\mathcal{N}|$ is the cardinality of $\mathcal{N}$. We call such an $\mathcal{N}$ a neighbourhood, and the elements of $\mathcal{N}$, neighbours of 0. Note that the symmetry and irreducibility of $\mathcal{N}$ imply

$$|\mathcal{N}| \geq 2d \text{ and is even for any neighbourhood } \mathcal{N}. \tag{1.2}$$

For $x \in \mathbb{Z}^d$, $x + \mathcal{N}$ is the set of neighbours of x. In the q-voter model the state at time t is $\xi_t : \mathbb{Z}^d \to \{0,1\}$, where $\xi_t(x)$ is the opinion of a voter at x at time t. The rate at which the voter at x changes opinion is the q-th power of the fraction of its neighbours with the opposite opinion. More formally, if $\xi \in \{0,1\}^{\mathbb{Z}^d}$ and

$$f_j(x,\xi) = \frac{1}{|\mathcal{N}|} \sum_{y \in \mathcal{N}} 1\{\xi(x+y) = j\}, \quad j = 0, 1,$$

and $c^{(q)}$ is defined by

$$c^{(q)}(x,\xi) = \widehat{\xi}(x) f_1^q(x,\xi) + \xi(x) f_0^q(x),$$

then the q-voter model ξ_t is the spin-flip process with rate function $c^{(q)}$ (see Theorem B.3 in [25]). Throughout we will use the notation

$$\widehat{\xi} = 1 - \xi,$$

and if $q = 0$, $0^q := 0$ in the above. The well-studied voter model is obtained by taking $q = 1$, in which case we will write $c^{\mathsf{vm}}(x,\xi)$ for $c^{(1)}(x,\xi)$.

The q-voter model was introduced by Nettle in [28] and also used by Abrams and Strogatz in [3] as a model of language death. The model, along with many variations of it, has been studied in the physics literature (e.g., see [1], [2], [22], [26], [30], [32]) Rigorous results for the model defined on large torii in $\mathbb{Z}^d$, $d \geq 3$ for q close to 1 have been obtained by Agarwal, Simper and Durrett in [4]. Our goal is to study the model on $\mathbb{Z}^2$ in the mathematically critical, and biologically important, two-dimensional case, with q close to 1 and $q < 1$. The methods we use to study this model will lead to some general results for a family of two-dimensional spin-flip processes.

We let $|\xi| = \sum_x \xi(x)$, and say that a probability measure ν on $\{0,1\}^{\mathbb{Z}^d}$ has the *coexistence property* if

$$\nu(|\widehat{\xi}| = |\xi| = \infty) = 1.$$

A translation invariant probability ν on $\{0,1\}^{\mathbb{Z}^d}$ has density $p \in [0,1]$ iff $\nu(\xi(0) = 1) = p$. In the case $d \geq 3$, it is well known that the voter model has a one-parameter family of translation invariant stationary distributions indexed by density, $\{\mu_\theta, 0 \leq \theta \leq 1\}$, such that for a wide class of initial laws with a given density, θ, the voter model converges in law as $t \to \infty$ to the corresponding μ_θ (see Chapter V of [24]). In particular, density is preserved over time. For $q < 1$, the flip rates to 1's say, are increased relative to the voter model by a factor of f_1^{q-1}, and so one is reinforcing the flip rates to locally rare types. That this effect strongly influences the ergodic behavior of the model is shown in Theorem 1.3 of [4], which considers a sequence of q_n-voter models $\xi_t^{(n)}$ on torii in $\mathbb{Z}^3$ of side length n, $q_n \uparrow 1$. That result shows that for an initial sequence of laws with any fixed density in (0,1), with an appropriate time rescaling, the density of $\xi_t^{(n)}$ approaches 1/2 and stays close to 1/2 for at least polynomially (in n) long times. We will refine this for $|\mathcal{N}| \leq 8$ (and in particular for the nearest neighbour setting in $d = 3, 4$) for $q < 1$, sufficiently close to one, by showing the existence of a translation invariant stationary distribution $\nu_{1/2}$ with the coexistence property and density $1/2$ to which the q-voter model converges weakly starting from any initial law with the coexistence property. See Theorem 1.1 where a slightly stronger "complete convergence" result is proved). In particular $\nu_{1/2}$ is the unique stationary law with the coexistence property (see Corollary 1.2).

Perhaps more interesting, such a convergence theorem also holds in two dimensions, again for $|\mathcal{N}| \leq 8$, and so, in particular, for the nearest neighbour case (see again Theorem 1.1). Recall that for $d = 2$, the voter model exhibits clustering, that is, it converges weakly to a mixture of all zeros and all ones as t tends to infinity. Dynamically, a typical site becomes part of a growing cluster of the same type as t becomes large. The actual type will change back and forth as time evolves and the clusters grow. A quantitative description of this dynamical clustering may be found in [8]. This clustering behaviour in two dimensions is typical of branching population models such as super-Brownian motion [15] or discrete time branching systems [23]. In these cases there is local extinction for large time and the mass becomes concentrated on larger and larger clumps separated by greater and greater distances. As the two-dimensional case is of particular importance in population modelling this failure to converge to any local equilibrium is referred to as the "pain in the torus". See [19] where a closely related model exhibiting the above clumping behaviour is dismissed as being "biologically irrelevant". The fact that, for any $q < 1$ and close to 1, the q-voter model converges to an essentially unique equilibrium is, we believe, of some general interest. Other interesting examples of such convergence results which include lower dimensions ($d \leq 2$), are due to Handjani [21] for the threshold voter model, corresponding to $q = 0$ (excluding only the one-dimensional nearest neighbour case where coexistence fails), the present authors [13] for the symmetric two-dimensional Lotka-Volterra model, and Sturm and Swart [31] for the one-dimensional "rebellious voter model" for sufficiently small competition parameter. Results and methods from the first two works will play a role here.

Our convergence results, and the coexistence results in [4], both require q close enough to 1. This is counter-intuitive because taking q smaller should only increase the advantage of locally rare types, described above. This restriction on q is due to the perturbative nature of both arguments which use the theory of voter model perturbations. This refers to the fact that the models approach the voter model as a parameter approaches a particular value–in this case as $q \uparrow 1$. [4] uses general results for voter model perturbations from [7] which require $d \geq 3$, as these results rely on rapid local convergence to the appropriate invariant law of the voter model. The two-dimensional case is more involved but, as was noted above, a particular class of voter model perturbations, Lotka-Volterra models, were analyzed (again using perturbative methods) for $d = 2$ in [12] and [9]. The methodologies in this latter paper will be extended here to show that rescaled q-voter models in which $q \uparrow 1$ as well, converge to a two-dimensional super-Brownian motion with a positive drift (see Section 1.5). The positivity of the drift will be critical to show that the process exhibits long-term coexistence. The description of the drift in terms of the asymptotics of long-time non-coalescing probabilities for two-dimensional random walks (see (1.22) in Section 1.3) is therefore important.

When $q > 1$ the above intuitive argument now goes in the other direction; the flip rates to 1's are multiplied by a factor of f_1^{q-1} and so, relatively speaking, one is reinforcing flip rates to locally dominant types. As a result we now expect a type of founder control, where one type or another will take over, with probabilities depending on the initial configuration. The tools for proving such a result for $q > 1$ and close enough to 1 in two dimensions, or even for $d \geq 3$, do not seem to be currently available.

In addition to being a voter model perturbation (see Section 3.2) the other key properties of the q-voter model used in the proof are *monotonicity* (see the definition in Section 1.2) and the *cancellative* property (see (3.1) in Section 3.1), which implies it has an *annihilating* dual. The fact that the q-voter model has no *coalescing* dual (see Section 4 of Chapter III in [24]) was established in [4], where it was shown that that the q-voter model ($q \neq 1$) is not additive.

Our "complete convergence" theorem for q-voter models is obtained as a corollary of a general result for two-dimensional monotone, cancellative finite range voter model perturbations when (the natural extension of) the above drift parameter is positive (see Theorem 1.9 below). This result will also imply such a "complete convergence" theorem with coexistence for (two-dimensional) affine voter models, geometric voter models and the aforementioned Lotka-Volterra models, for appropriate choices of parameter (see the examples in Section 3.2 and then Theorems 5.6, 5.7 and 5.8). We note that no restrictions on $|\mathcal{N}|$ are required for these three models or for Theorem 1.9 to hold. The Lotka-Volterra result was first proved in [13], while the other applications are new.

1.2 A complete convergence theorem for the q-voter model

Assume first $c(x,\xi)$ is a rate function satisfying condition (B4) of [25] (see (4.3) below), and so by Theorem B3 of that reference is the rate function of a unique spin flip system ξ_t starting in state ξ_0 under P_{ξ_0}. Define the hitting times $\tau_{\mathbf{0}} = \inf\{t \geq 0 : \xi_t = \mathbf{0}\}$ and $\tau_{\mathbf{1}} = \inf\{t \geq 0 : \xi_t = \mathbf{1}\}$. We identify a random vector with its probability law, as usual, and introduce the probabilities

$$\beta_0(\xi_0) = P_{\xi_0}(\tau_{\mathbf{0}} < \infty), \quad \beta_1(\xi_0) = P_{\xi_0}(\tau_{\mathbf{1}} < \infty), \quad \beta_\infty(\xi_0) = P_{\xi_0}(\tau_{\mathbf{0}} = \tau_{\mathbf{1}} = \infty).$$

By standard arguments (see (1.8) of [13]),

$$\beta_0(\xi_0) = 0 \text{ if } |\xi_0| = \infty, \; \beta_1(\xi_0) = 0 \text{ if } |\widehat{\xi_0}| = \infty, \text{ and hence } \beta_\infty(\xi_0) = 1 \text{ if } |\xi_0| = |\widehat{\xi_0}| = \infty. \tag{1.3}$$

Recall that a spin-flip process with rate function $c(x,\xi)$ is monotone iff for every $\underline{\xi} \leq \xi$ (this means $\underline{\xi}(x) \leq \xi(x)$ for all x),

$$\begin{aligned} c(x,\xi) &\geq c(x,\underline{\xi}) \text{ if } \xi(x) = \underline{\xi}(x) = 0, \\ c(x,\xi) &\leq c(x,\underline{\xi}) \text{ if } \xi(x) = \underline{\xi}(x) = 1, \end{aligned}$$

We assume throughout that $0 \leq q \leq 1$ and note that clearly, for all such q,

$$\text{the } q\text{-voter model is monotone.} \tag{1.4}$$

We let $\Rightarrow$ denote weak convergence of probability laws. A probability ν on $\{0,1\}^{\mathbb{Z}^d}$ is symmetric iff $\nu(\widehat{\xi} \in \cdot) = \nu(\xi \in \cdot)$. Both $\mathbf{0}$ and $\mathbf{1}$ (the configurations of all 0's and all 1's, respectively) are obviously traps for the q-voter model.

Theorem 1.1. *Assume $|\mathcal{N}| \leq 8$ and $d = 2, 3$ or 4. There exists $0 < q_c < 1$ such that for $q_c < q < 1$ there is a translation invariant symmetric stationary distribution $\nu_{1/2}$ with density $1/2$ satisfying the coexistence property, and such that for all initial $\xi_0 \in \{0,1\}^{\mathbb{Z}^d}$,*

$$\xi_t \Rightarrow \beta_0(\xi_0)\delta_{\mathbf{0}} + \beta_\infty(\xi_0)\nu_{1/2} + \beta_1(\xi_0)\delta_{\mathbf{1}} \quad \text{as } t \to \infty, \tag{1.5}$$

where $\beta_\infty(\xi_0) > 0$ unless ξ_0 is one of the fixed points, $\mathbf{0}$ or $\mathbf{1}$.

We will state Theorem 1.1 more compactly by saying that for $q_c < q < 1$ *the complete convergence theorem with coexistence (CCT) holds for the q-voter model.* As an immediate corollary we have a complete description of all invariant laws.

Corollary 1.2. *For d, $\mathcal{N}$, and q_c as in Theorem 1.1, and $q_c < q < 1$, $\nu_{1/2}$ is the only invariant distribution with the coexistence property, and $\{\nu_{1/2}, \delta_{\mathbf{0}}, \delta_{\mathbf{1}}\}$ are the only extremal invariant distributions.*

Remark 1.3. Note that the hypotheses of the above results are satisfied when $\mathcal{N}$ is the set of nearest neighbours in $\mathbb{Z}^d$ and $2 \le d \le 4$. They also hold for $d = 2$ when $\mathcal{N}$ is the unit sphere in $\mathbb{Z}^2$ in the L^∞ norm ($|\mathcal{N}| = 8$). Clearly the restriction $|\mathcal{N}| \le 8$ forces $d \le 4$ by (1.2).

The q-voter model is a *nonlinear voter model* as defined in [5]. These are spin-flip processes where the rate of flipping depends only on the number of sites of the opposite type in $\mathcal{N}$. The cancellative property is defined in Section III.4 of [24] and discussed in Section 3.1 below. It is verified for the q-voter model in Lemma 3.3 when $|\mathcal{N}| \le 8$ and $q \in (q_c, 1)$ using a criteria from [5] developed for nonlinear voter models (Proposition 3.2). This criteria involves the inverse of an $|\mathcal{N}| \times |\mathcal{N}|$ matrix and so becomes more complicated as $|\mathcal{N}|$ increases. This is the reason we restrict the neighbourhood size in Theorem 1.1.

Conjecture 1.4. *In Theorem 1.1 complete convergence with coexistence continues to hold for any neighbourhood $\mathcal{N}$ in any dimension $d \ge 2$ for $q < 1$ and sufficiently close to 1.*

As the above discussion suggests, and a more careful analysis of the proofs shows, this would follow from Conjecture 3.4 on the cancellative property holding for general q-voter models. We hasten to add that for other models like the geometric voter model, the threshold voter, and the affine voter model, one can apply Proposition 3.2 for arbitrary $\mathcal{N}$, and for the Lotka-Volterra model, one can check the cancellative property directly through an educated guess of the annihilating dual (see [29] and Section 6 of [13]). These models are defined in Section 3.2.

As already noted in Section 1.1, taking $q < 1$ smaller than q_c should only make large clusters less likely and so the complete convergence should follow for all $0 < q < 1$ for $d \ge 2$. Moreover, as the result fails for $q = 1$ (the voter model), we expect q close to 1, handled to some extent in Theorem 1.1, to be the most delicate case. Recall also from Section 1.1 that for the extreme case $q = 0$, a CCT is proved in [21] for $d \ge 2$.

Conjecture 1.5. *In Theorem 1.1 complete convergence with coexistence holds for any neighbourhood $\mathcal{N}$, any $d \ge 2$ and any $0 < q < 1$.*

We state the CCT for $0 < q < 1$ and $d \ge 2$ as a separate conjecture because we believe the issues here are quite different from those underlying Conjecture 3.4. Here a result in any d for any neighbourhood would be of great interest. Similar “obvious" results should also hold for the CCT for Lotka-Volterra models for (symmetric) competition parameter $\alpha \in (0, 1)$ but are again only proved for the most delicate case when α near 1 (see Theorem 1.1 of [13]) due to the perturbative nature of the proofs.

To state our general two-dimensional CCT we need to introduce the drift parameter mentioned in Section 1.1, and for this, we first need some long time asymptotics of non-coalescing probabilities for two-dimensional random walks.

1.3 Two-dimensional coalescing random walk

Let $p : \mathbb{Z}^2 \to [0, 1]$ be a symmetric, irreducible, random walk kernel with covariance matrix $\sigma^2 I$ for some $\sigma > 0$, such that $p(0) = 0$. For the particular case of q-voter models, $p(\cdot)$ will be the uniform law on a neighbourhood $\mathcal{N}$. Under a probability $\hat{P}$, let $\{B_t^x, x \in \mathbb{Z}^2\}$ be a system of rate one continuous time coalescing random walks with jump kernel p, and for $A \subset \mathbb{Z}^2$ define $B_t^A = \{B_t^x, x \in A\}$, let $|B_t^A|$ denote its cardinality, and let the time it takes all walks starting in A to coalesce to a single walk be $\tau(A) = \inf\{t \geq 0 : |B_t^A| = 1\}$. For $n \geq 2$ and nonempty, finite disjoint $A_1, \dots, A_n \subset \mathbb{Z}^2$, define the stopping times

$$\begin{aligned}&\tau(A_1, \dots, A_n) = \max_{1 \leq i \leq n} \tau(A_i), \\ &\sigma(A_1, \dots, A_n) = \inf\{t \geq 0 : B_t^{A_i} \cap B_t^{A_j} \neq \emptyset \text{ for some } i \neq j\}.\end{aligned} \tag{1.6}$$

At the risk of some confusion, for $x \in \mathbb{Z}^2$, we will often identify $\{x\}$ with x. We write

$$a(t) \sim b(t) \text{ as } t \to \infty \text{ to mean } \lim_{t \to \infty} \frac{a(t)}{b(t)} = 1.$$

Proposition 1.3 in [9] states that for $n \geq 2$ and distinct $x_1, \dots, x_n \in \mathbb{Z}^d$ there is a finite $K_n(x_1, \dots, x_n) > 0$ such that

$$q_{x_1,\dots,x_n}(t) := \hat{P}\big(\sigma(x_1, \dots, x_n) > t\big) \sim \frac{K_n(x_1, \dots, x_n)}{(\log t)^{\binom{n}{2}}} \text{ as } t \to \infty. \tag{1.7}$$

The $n = 2$ case is well known, as $\sigma(x_1, x_2)$ has the same law as the hitting time of 0 of a walk starting at $x_1 - x_2$ run at rate 2. The following extension of (1.7), proved in Section 2 below, will be used to describe the key drift term arising in our general CCT.

Proposition 1.6. *Let $n \geq 2$ and $A_1, \dots, A_n$ be nonempty finite disjoint subsets of $\mathbb{Z}^2$. Then there exists a finite $K_n(A_1, \dots, A_n) > 0$ such that*

$$\hat{P}(\sigma(A_1, \dots, A_n) > t, \tau(A_1, \dots, A_n) < t) \sim \frac{K_n(A_1, \dots, A_n)}{(\log t)^{\binom{n}{2}}} \text{ as } t \to \infty. \tag{1.8}$$

In fact, if $a_i \in A_i$, $1 \leq i \leq n$,

$$\begin{aligned}&K_n(A_1, \dots, A_n) \\ &\quad = \sum_{\text{distinct } x_1,\dots,x_n \in \mathbb{Z}^2} K_n(x_1, \dots, x_n) \hat{P}(\sigma(A_1, \dots, A_n) > \tau(A_1, \dots, A_n), \\ &\qquad\qquad B^{a_i}_{\tau(A_1,\dots,A_n)} = x_i,\ 1 \leq i \leq n).\end{aligned} \tag{1.9}$$

Remark 1.7. Let S be a finite subset of $\mathbb{Z}^2$. By summing over partitions of S of cardinality n, one sees that for $n \geq 2$, if $u(t) \geq Ct^r$ for some $C, r > 0$, then

$$\sup_{t \geq 1} (\log t)^{\binom{n}{2}} \hat{P}(|B^S_{u(t)}| = n) < \infty.$$

Let $\mathcal{N}$ be a neighbourhood in $\mathbb{Z}^2$ (in practice it will contain the support of p but this is not needed for our definitions), and set

$$\bar{\mathcal{N}} = \mathcal{N} \cup \{0\}.$$

For a set Γ, $|\Gamma| \geq k$, let $\mathcal{P}_k(\Gamma)$ be the set of partitions $\{\pi_1, \dots, \pi_k\}$ of Γ such that each $|\pi_i| \geq 1$. We will write $\mathcal{P}(\Gamma)$ for $\mathcal{P}_2(\Gamma)$. For A a non-empty subset of $\mathcal{N}$, define

$$\begin{aligned}\Theta^+(A) &= \sum_{\{A_1,A_2\} \in \mathcal{P}(\bar{\mathcal{N}} \setminus A)} K_3(A, A_1, A_2), \\ \Theta^-(A) &= \sum_{\{A_1,A_2\} \in \mathcal{P}(A)} K_3(\bar{\mathcal{N}} \setminus A, A_1, A_2).\end{aligned} \tag{1.10}$$

Note that we have suppressed the dependence of $\Theta^{\pm}$ on p and $\mathcal{N}$. For the q-voter model it will be understood that $p = 1_{\mathcal{N}}/|\mathcal{N}|$ for a given neighbourhood $\mathcal{N}$.

1.4 A general complete convergence theorem in two dimensions

The main conditions we impose on our spin-flip system is that they are cancellative and constitute a finite range voter model perturbation. To define the latter, for $d \geq 2$, let $p : \mathbb{Z}^d \to [0,1]$ be a symmetric, irreducible, random walk kernel with finite support and $p(0) = 0$ (as in the last section but now with $d \geq 2$). Write $p(A)$ for $\sum_{y \in A} p(y)$. Assume that

$$p \text{ has covariance matrix } \sigma^2 I \text{ for some } \sigma > 0. \tag{1.11}$$

Let $f_i(x,\xi) = \sum_y p(y-x) 1\{\xi(y) = i\}$ (agreeing with our earlier notation if p is uniform on $\mathcal{N}$) and introduce the associated voter model rates

$$c^{\mathsf{vm}}(x,\xi) = \widehat{\xi}(x) f_1(x,\xi) + \xi(x) f_0(x,\xi).$$

Consider also a neighbourhood $\mathcal{N}$ containing the support of p. We write $\xi|_{x+\mathcal{N}}$ for the function on $\mathcal{N}$ which maps $y \in \mathcal{N}$ to $\xi(x+y)$.

Definition 1.8. *A voter model perturbation on $\mathbb{Z}^d$ for $d \geq 2$ with finite range in $\mathcal{N}$ is a family of translation invariant spin-flip systems, $\{\xi^{[\varepsilon]}_\cdot : 0 < \varepsilon \leq \varepsilon_0\}$, for some $\varepsilon_0 \in (0,1]$, with rate functions*

$$c_\varepsilon(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \varepsilon c^*_\varepsilon(x,\xi) \geq 0 \quad \text{for all } x \in \mathbb{Z}^d,\ \xi \in \{0,1\}^{\mathbb{Z}^d}, \tag{1.12}$$

where for some $g_0^\varepsilon, g_1^\varepsilon : \{0,1\}^{\mathcal{N}} \to \mathbb{R}$,

$$c^*_\varepsilon(x,\xi) = \widehat{\xi}(x) g_1^\varepsilon(\xi|_{x+\mathcal{N}}) + \xi(x) g_0^\varepsilon(\xi|_{x+\mathcal{N}}). \tag{1.13}$$

In addition there are $g_i : \{0,1\}^{\mathcal{N}} \to \mathbb{R}$ such that

$$\|g_i^\varepsilon - g_i\|_\infty \leq c_g \varepsilon^{r_0} \text{ for } i = 0,1 \text{ and all } \varepsilon \text{ and some } c_g, r_0 > 0, \text{ if } d \geq 3, \tag{1.14}$$

and

$$\lim_{\varepsilon \to 0+} \|g_i^\varepsilon - g_i\|_\infty = 0 \text{ for } i = 0,1, \text{ if } d = 2. \tag{1.15}$$

Finally we assume

$$\text{for all } \varepsilon \in (0,\varepsilon_0],\ \mathbf{0} \text{ is a trap for } \xi^{[\varepsilon]},\ \text{that is, } g_1^\varepsilon(1_\emptyset) = 0, \tag{1.16}$$

and, in addition if $d = 2$,

$$\text{for all } \varepsilon \in (0,\varepsilon_0],\ \mathbf{1} \text{ is a trap for } \xi^{[\varepsilon]},\ \text{that is, } g_0^\varepsilon(1_{\mathcal{N}}) = 0. \quad \square \tag{1.17}$$

This class of processes is discussed further in Section 3.2. At times we will abuse the wording and say $\xi^{[\varepsilon]}$ is a finite range voter model perturbation, for $0 < \varepsilon \leq \varepsilon_0$. The following "asymptotic rate function" associated with the above finite range voter perturbation will play an important role:

$$r^s(A) := g_1(1_A) = \lim_{\varepsilon \to 0} \frac{c_\varepsilon(0,1_A) - f_1(0,1_A)}{\varepsilon} \quad \text{for } A \subset \mathcal{N}. \tag{1.18}$$

The above equality is elementary. For $d = 2$ the "drift" associated with the above voter model perturbation with finite range in $\mathcal{N}$ is

$$\Theta_3 := \sum_{\emptyset \neq A \subset \mathcal{N}} r^s(A)(\Theta^+(A) - \Theta^-(A)). \tag{1.19}$$

Fix a neighbourhood $\mathcal{N}$ in $\mathbb{Z}^d$ where $d \geq 2$. It is easy to see the family of associated q-voter models for $q = 1 - \varepsilon$ is a finite range voter model perturbation (see Example 3.8 in Section 3.2). If

$$r_\ell = (\ell/|\mathcal{N}|)\log(|\mathcal{N}|/\ell), \text{ for } \ell = 1,\ldots,|\mathcal{N}|, \text{ and } r_0 = 0, \tag{1.20}$$

then for the family of q voter models and for $A \subset \mathcal{N}$,

$$r^s(A) = \lim_{\varepsilon \to 0} \frac{(|A|/|\mathcal{N}|)^{1-\varepsilon} - (|A|/|\mathcal{N}|)}{\varepsilon} = r_{|A|}. \tag{1.21}$$

Therefore, for the two-dimensional q-voter model we have

$$\Theta_3 := \Theta = \sum_{\emptyset \neq A \subset \mathcal{N}} r_{|A|}(\Theta^+(A) - \Theta^-(A)). \tag{1.22}$$

Here is our general complete convergence theorem in two dimensions.

Theorem 1.9. *Assume for $0 < \varepsilon \le \varepsilon_0$, $\xi^{[\varepsilon]}$ is a cancellative and monotone finite range voter model perturbation in $\mathbb{Z}^2$, and $\Theta_3 > 0$. There is an $\varepsilon_1 > 0$ such that for $\varepsilon \in (0, \varepsilon_1)$, the complete convergence theorem with coexistence (CCT) holds for $\xi^{[\varepsilon]}$.*

Turning to the q-voter model in two dimensions with $|\mathcal{N}| \le 8$, we have already noted that in any dimension this model is monotone (elementary), cancellative (Lemma 3.3) and a finite range voter model perturbation (Example 3.8). So it remains to verify that $\Theta_3 > 0$, which clearly is a crucial condition, as all the other conditions hold equally well for the ordinary voter model where the CCT fails, and $\Theta_3 = 0$. In Corollary 5.3 we will show in complete generality that $\Theta_3 > 0$ will easily follow from the strict subadditivity of r^s, that is from

$$r^s(A \cup B) < r^s(A) + r^s(B) \text{ for all non-empty disjoint } A, B \subset \mathcal{N}. \tag{1.23}$$

For the q-voter model this means (recall (1.21)) if r_ℓ is as in (1.20), then

$$r_{\ell_1+\ell_2} < r_{\ell_1} + r_{\ell_2} \text{ for all } 0 < \ell_i,\ \ell_1 + \ell_2 \le |\mathcal{N}|. \tag{1.24}$$

This follows from an elementary calculus exercise, and was noted in Section 5 of [4], where it played an important role in their analysis of the q-voter model for $d \ge 3$. The general condition (1.23) owes much to the calculation in [4]. Therefore Corollary 5.3 and (1.24) imply that

$$\Theta > 0. \tag{1.25}$$

We now may apply this and Theorem 1.9 to prove Theorem 1.1 for $d = 2$. See Corollary 5.5 for a general statement of this reasoning to establish a CCT in two dimensions.

Theorem 1.9 is a two-dimensional version of Theorem 1.2 of [13] where a similar result is stated for $d \ge 3$. More specifically for $d \ge 3$ this result establishes a complete convergence theorem with coexistence for cancellative spin-flip systems which are voter model perturbations (as defined in Section 1 of [13]), providing a certain drift is positive. The drift is $f'(0)$ where

$$\frac{\partial u}{\partial t} = \sigma^2 \frac{\Delta u}{2} + f(u),$$

is the limiting reaction diffusion equation under law of large numbers scaling (see [4] or [7]). In fact the drift, $f'(0)$, also equals the positive drift in a limiting super Brownian motion arising in a low density scaling theorem (see Corollary 1.8 of [10]). That these two drifts coincide is easy and shown on pages 33-34 in Section 1.8 of [7], and the fact that the hypotheses of Corollary 1.8 of [10] hold for voter model perturbations is also verified in the same place. In either representation, the positive drift is used to show regions of low density will repopulate to avoid local extinction and the resulting clumping. The reaction function f is defined in terms of the invariant measures for the voter model and so is not well-defined for $d = 2$. Therefore, in extending Theorem 1.2 of [13] to two dimensions in Theorem 1.9 we replace $f'(0)$ with the drift Θ_3 in a super Brownian motion low density limit theorem (Theorem 1.15 and Remark 1.16 in the next section). See Section 1.5 for more about this convergence to super-Brownian motion.

Remark 1.10. In Theorem 1.2 of [13] the last condition for a (CCT), namely $\beta_\infty(\xi^{[\varepsilon]}) > 0$ if $\xi^{[\varepsilon]}$ is not $\mathbf{0}$ or $\mathbf{1}$, was not part of the conclusion, but in fact it is easy to argue just as in the 2-dimensional result above to derive this condition. See Remark 4.7 below.

Consider briefly the simpler $d \geq 3$ case of Theorem 1.1. We have noted that the hypotheses of Theorem 1.2 of [13] have been verified, at least for $|\mathcal{N}| \leq 8$, aside from the positivity of $f'(0)$. This last property follows from Theorem 1.2 of [4] (the proof given there for $d = 3$ holds in any dimension). As a result, we are then able to establish the $d \geq 3$ case of Theorem 1.1 as a direct consequence of Theorem 1.2 of [13]. This simple argument is carried out in Section 3.3 below. The same lemma (Lemma 5.2) which led to the positivity of Θ_3 under the strict subadditivity of r^s in Corollary 5.3, also leads to a simple self-contained direct proof of $f'(0) > 0$ for $d \geq 3$ (see Proposition 5.9).

To prove the 2-dimensional result, Theorem 1.9, we will use the more fundamental Proposition 4.1 of [13], instead of Theorem 1.2 in [13]. The former result holds for arbitrary d at the cost of bringing in some additional technical hypotheses. (In fact this result was used to establish Theorem 2.1 in [13] for $d \geq 3$.) Proposition 4.1 of [13] is combined with several other results in [13] to prove Theorem 4.4 below. This result establishes the CCT for cancellative finite range voter model perturbations if two additional conditions ((4.10) and (4.11)) are in force. These additional conditions demonstrate the ability of 1's and 0's to both coexist and propagate in space and time, respectively. The next step is to follow the derivation of the CCT for the two-dimensional Lotka-Volterra model in Section 6 of [13] to show that the above conditions will follow from a set-up which allows a comparison to super-critical oriented percolation (see (4.18) below) to show simultaneous propagation of both 0's and 1's in close proximity as time gets large. This is done in Theorem 4.6. The last step is to justify the above super-critical oriented percolation set-up by using a low density limit theorem in which the scaling limit is a super-Brownian motion with *positive drift* Θ_3 (Theorem 1.15 and Remark 1.16 in the next section.)

1.5 A scaling limit theorem in 2 dimensions

First consider a scaling limit theorem for the q-voter model. We will speed up time and scale down space for a two-dimensional q-voter model in the usual Brownian manner and at the same time let $q \uparrow 1$ at an appropriate rate. In the regime where 1's are relatively rare we show the normalized empirical measure of 1's converges to super-Brownian motion with drift. Such limit theorems are technically more difficult in two dimensions than higher dimensions, even in the simple voter model setting [6]. The increased clustering in two dimensions (e.g from the stronger recurrence of the dual for the voter model) leads to a greater branching rate, or equivalently, a greater mass per particle. The resulting extra $\log N$ factor complicates even the simplest moment bounds. Convergence to super-Brownian motion in two dimensions was established for a class of Lotka-Volterra spin systems (also voter model perturbations) in Theorem 1.5 of [9]. We will refine some of the results and methods used in that paper to obtain the required scaling limit theorem. In [9] a particular branching coalescing dual process was used in some key calculations. Such duals seem more complex in our present setting and so instead we use a systematic comparison of our model with the voter model over small intervals (see, for example, Section 7.1). We believe this gives a more robust approach to general voter model perturbations in the critical and physically important two-dimensional case. The higher dimensional ($d \geq 3$) analogues of this limit result follow from [10] which proves a limit theorem for a general class of voter model perturbations including q-voter models as $q \uparrow 1$ (see Remark 1.13 below).

Let $\mathcal{M}_F = \mathcal{M}_F(\mathbb{R}^2)$ denote the space of finite measures on $\mathbb{R}^2$ with the topology of weak convergence. A 2-dimensional super-Brownian motion with initial condition $X_0 \in \mathcal{M}_F(\mathbb{R}^2)$, branching rate $b > 0$, diffusion coefficient $\sigma^2 > 0$, and drift $\theta \in \mathbb{R}$, denoted $SBM(X_0, b, \sigma^2, \theta)$, is an $\mathcal{M}_F(\mathbb{R}^2)$-valued diffusion X whose law is the unique solution of the martingale problem:

$$\text{(MP)} \quad \begin{cases} \forall\, \phi \in C_b^3(\mathbb{R}^2), \quad M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s \left(\frac{\sigma^2}{2}\Delta\phi + \theta\phi\right) ds \\ \quad \text{is a continuous } \mathcal{F}_t^X\text{-martingale such that} \\ \quad \langle M(\phi)\rangle_t = \int_0^t X_s \left(b\phi^2\right) ds. \end{cases}$$

Here C_b^3 is the set of bounded C^3 functions with bounded continuous partials of order 3 or less and $\mathcal{F}_t^X$ is the canonical right-continuous filtration generated by X.

For $N \geq e^3$ (N denotes a real number), let

$$N' = N/\log N, \text{ and } \varepsilon_N = (\log N)^3/N. \tag{1.26}$$

We let $\xi_t^{(q_N)}$ denote a q_N-voter model on $\mathbb{Z}^2$ with $q_N = 1 - \varepsilon_N$. Consider the rescaled q_N-voter model,

$$\xi_t^N(x) = \xi_{Nt}^{(q_N)}(x\sqrt{N}), \quad x \in S_N := \mathbb{Z}^2/\sqrt{N},$$

and define the associated $\mathcal{M}_F$-valued empirical process

$$X_t^N = \frac{1}{N'} \sum_{x \in S_N} \xi_t^N(x)\delta_x. \tag{1.27}$$

Theorem 1.11. *Assume $\mathcal{N}$ is a neighbourhood in $\mathbb{Z}^2$, $\sigma^2 = \sigma^2(\mathcal{N})$ is as in* (1.1) *and $\{\xi_0^N\}$ satisfies $X_0^N \to X_0$ in $\mathcal{M}_F$. If Θ is as in* (1.22)*, then $\Theta > 0$ and*

$$X^N \Rightarrow \textit{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \Theta) \textit{ in the Skorokhod space } D(\mathbb{R}_+, \mathcal{M}_F) \textit{ as } N \to \infty.$$

The fact that $\Theta > 0$ was already noted above in (1.25).

Remark 1.12. There is a symmetric result for $q > 1$ where we take $q_N = 1 + \varepsilon_N$. As noted in [4], the r_ℓ in this case is the negative of the r_ℓ in (1.20) and therefore in (1.22), $\Theta_3 < 0$ will have the opposite sign. The same proof noted below in Remark 1.16 (using the more general Theorem 1.15) then gives the conclusion of Theorem 1.11 with drift $\Theta_3 < 0$, the negative of that in Theorem 1.11.

Remark 1.13. The analogue of Theorem 1.11 for $d \geq 3$ follows from a limit theorem for a class of voter model perturbations established as Corollary 1.8 of [10]. In this setting we take $N' = N$ and $\varepsilon_N = 1/N$ in our definition of X^N. It is then straightforward to verify the hypotheses of the above result and so conclude that for $\mathcal{N}$, σ^2, and $\{X_0^N\}$ as in Theorem 1.11 with $d \geq 3$,

$$X^N \Rightarrow \text{SBM}(X_0, 2\gamma_e, \sigma^2, \Theta) \text{ in } D(\mathbb{R}_+, \mathcal{M}_F(\mathbb{R}^d)) \text{ as } N \to \infty.$$

Here $\gamma_e \in (0,1)$ is the escape probability from 0 of the random walk in $\mathbb{Z}^d$ whose step kernel is uniform in $\mathcal{N}$, and

$$\Theta = \sum_{\emptyset \neq A \subset \mathcal{N}} \beta(A)\hat{P}(\tau(A) < \infty, \tau(A \cup \{0\}) = \infty) - \delta(A)\hat{P}(\tau(A \cup \{0\}) < \infty) > 0,$$

where for r_ℓ as in (1.20),

$$\beta(A) = \sum_{\emptyset \neq C \subset A} 1(C \neq \mathcal{N})(-1)^{|A|-|C|} r_{|C|}, \text{ and } \delta(A) = \sum_{\emptyset \neq C \subset A} 1(C \neq \mathcal{N})(-1)^{|A|-|C|} r_{|\mathcal{N} \setminus C|}.$$

The positivity of Θ follows from Theorem 1.2 of [4] (or Proposition 5.9 below) and the fact that the drift Θ agrees with $f'(0)$ where f is the reaction function in the limiting reaction diffusion equation (see Section 1.8 of [7]), denoted by ϕ in [4].

We now consider a general two-dimensional limit theorem for a large class of finite range voter model perturbations. The random walk kernel, p, is as described at the start of Section 1.4 with $d = 2$. The key condition is the following “asymptotic $0-1$ symmetry" which was used implicitly for the special case of two-dimensional Lotka-Volterrra models in [9].

Definition 1.14. *Consider a finite range voter model perturbation in $\mathbb{Z}^2$, $\xi^{[\varepsilon]}$, as in Definition 1.8 with rates $c_\varepsilon(x,\xi)$ for $0 < \varepsilon \le \varepsilon_0$. We say $\xi^{[\varepsilon]}$ is asymptotically symmetric if, in addition, for some $g^a : \{0,1\}^{\mathcal{N}} \to \mathbb{R}$,*

$$\lim_{\varepsilon\to 0+}(\log 1/\varepsilon)^2(g_1^\varepsilon(\xi) - g_0^\varepsilon(\widehat{\xi})) = g^a(\xi) \text{ for all } \xi \in \{0,1\}^{\mathcal{N}}, \tag{1.28}$$

or, equivalently, for some $c^a : \mathbb{Z}^2 \times \{0,1\}^{\mathbb{Z}^2} \to \mathbb{R}$ (necessarily anti-symmetric),

$$\lim_{\varepsilon\to 0+}(\log 1/\varepsilon)^2 \frac{c_\varepsilon(x,\xi) - c_\varepsilon(x,\widehat{\xi})}{\varepsilon} = c^a(x,\xi) \text{ for all } x \in \mathbb{Z}^2 \text{ and } \xi \in \{0,1\}^{\mathbb{Z}^2}. \tag{1.29}$$

The relationship between c^a and g^a is (note that c^a is translation invariant)

$$c^a(0,\xi) = \widehat{\xi}(0)g^a(\xi|_{\mathcal{N}}) - \xi(0)g^a(\widehat{\xi}|_{\mathcal{N}}).$$

Further discussion may be found in Section 3.2. The following asymmetric function will also be important for our limit theorem:

$$r^a(A) := g^a(1_A) = \lim_{\varepsilon\to 0}(\log(1/\varepsilon))^2 \frac{c_\varepsilon(0,1_A) - c_\varepsilon(0,1_{\bar{\mathcal{N}}\setminus A})}{\varepsilon}. \tag{1.30}$$

The above equality is again elementary.

We are ready to state our general two-dimensional limit theorem. It will not require monotonicity or the cancellative property. Recall the notation $K_2(A_1,A_2)$ from Section 1.3, and that σ^2 is as in (1.11). A second “drift" parameter associated with such voter model perturbations will be denoted by

$$\Theta_2 = \sum_{\emptyset\neq A\subset\mathcal{N}} r^a(A)K_2(A,\bar{\mathcal{N}}\setminus A). \tag{1.31}$$

Theorem 1.15. *Assume $\{\xi^{[\varepsilon]} : 0 < \varepsilon \le \varepsilon_0\}$ is an asymptotically symmetric finite range voter model perturbation on $\mathbb{Z}^2$. Let $\xi_t^N(x) = \xi_{Nt}^{[\varepsilon_N]}(x\sqrt{N})$, $x \in S_N$. Define a measure-valued process by*

$$X_t^N = (1/N')\sum_{x\in S_N}\xi_t^N(x)\delta_x. \tag{1.32}$$

If $X_0^N \to X_0$ in $\mathcal{M}_F$, then

$$X^N \Rightarrow \text{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \Theta_2+\Theta_3) \text{ in the Skorokhod space } D(\mathbb{R}_+,\mathcal{M}_F) \text{ as } N\to\infty.$$

Remark 1.16. Clearly a finite range voter model perturbation which is symmetric, that is $c_\varepsilon(x,\xi) = c_\varepsilon(x,\widehat{\xi})$ for all x and ξ, is asymptotically symmetric with $c^a = r^a = \Theta_2 = 0$. It will turn out that a cancellative finite range voter model perturbation is necessarily symmetric for each ε (see Remark 4.5). Therefore in the setting of our general CCT (Theorem 1.9), the above limit theorem applies with $\Theta_2 = 0$ and the drift for our limiting SBM is indeed Θ_3, as in the discussion in Section 1.4. In particular, this is the case for the q-voter model. Therefore, recalling (1.22) and (1.25), we see that Theorem 1.11 is an immediate consequence of the general Theorem 1.15 above.

Theorem 1.15 includes all the examples we know of super-Brownian limits for voter model perturbations in two dimensions, as well as a number of new ones. In addition to the above result for q-voter models and the limit theorem for the ordinary 2-dimensional voter model in [6], this includes the basic limit theorems for Lotka-Volterra models in [12] (see Example 6.6), the more refined Lotka-Volterra limit theorems in [9] (see Example 6.2), and limit theorems for the affine and geometric voter model (Examples 6.3 and 6.4, respectively).

The following “survival" corollary is an easy consequence of Theorem 1.15 and standard arguments (see Section 10).

Corollary 1.17. *Assume for* $0 < \varepsilon \le \varepsilon_0$, $\xi^{[\varepsilon]}$ *is a monotone asymptotically symmetric finite range voter model perturbation in* $\mathbb{Z}^2$, *and* $\Theta_2 + \Theta_3 > 0$. *There is an* $\varepsilon_1 > 0$ *such that for* $\varepsilon \in (0, \varepsilon_1)$, $P_{\delta_0}(|\xi_t^{[\varepsilon]}| > 0$ *for all* $t \ge 0) > 0$.

Section 2 discusses coalescing random walks and proves Proposition 1.6. Cancellative processes are defined, and many of their properties are presented, in Section 3.1. Here the criterion for a nonlinear voter model to be cancellative (Proposition 3.2 from [5]) is proved for completeness, and then applied to show the q-voter model is cancellative for all $q \in [0,1]$ if $|\mathcal{N}| = 4$. The $|\mathcal{N}| = 8$ case for q near 1 and $q < 1$ is outlined here, while the actual `maple`-assisted proof is presented in an Appendix. Additional properties of finite range voter model perturbations and associated notation are presented in Section 3.2. Several examples of cancellative finite range voter model perturbations are presented here as well. The short proof of Theorem 1.1 for $d \ge 3$ is presented in Section 3.3. The general complete convergence theorem, Theorem 1.9, is proved in Section 4, assuming Theorem 4.9, which states that the conditions for a block comparison to super-critical percolation, (4.18), will hold for a monotone, asymptotically symmetric, finite range voter model perturbation, if $\Theta_2 + \Theta_3 > 0$. This section also includes a coupled SDE construction of our particle system ξ along with $\widehat{\xi}$, and killed versions of these processes, on a common probability space (Proposition 4.3 and the ensuing (4.9)). This set-up is used in the proof of an intermediate result (Theorem 4.6) in which a CCT is established assuming (4.18) in place of $\Theta_2 + \Theta_3 > 0$. Theorem 4.9 is proved in Section 10 as a corollary to the weak convergence result, Theorem 1.15, *and its proof*. The latter result is proved in Sections 6-8 and Section 9. Section 6 sets up the approximating martingale problems, using a convenient SDE coupling, and also gives a number of examples of the weak convergence theorem. A number of preliminary bounds and sharp estimates on the drift terms are given in Section 7, while the proof of convergence to SBM is in Section 8. A key technical bound, Proposition 7.14, used for exact asymptotics on the drift terms, is proved in Section 9.

Acknowledgement. It is a pleasure to thank Mathieu Merle for his help with the proof of Proposition 1.6.

2 Coalescing probability asymptotics for two-dimensional random walks

We work in the setting of Section 1.3, and in particular, $p(\cdot)$ is the general random walk kernel considered there, and $q_{x_1,\dots,x_n}(t)$ and $K_n(x_1,\dots,x_n)$ are as in (1.7). In fact, as the reader can easily check, the proof below includes a derivation of (1.7).

Lemma 2.1. *Assume* $n \ge 2$ *and* $x_1, \dots, x_n \in \mathbb{Z}^2$ *are distinct. There are positive constants* $C_{2.1}$ *and* $t_0 > 0$ *depending only on* p *and* n *such that*

$$(\log t)^{\binom{n}{2}} q_{x_1,\dots,x_n}(t) \le C_{2.1} K_n(x_1,\dots,x_n) \text{ if } t \ge 2\big(\max_{i \ne j}\{|x_i - x_j|^4\} \vee t_0\big). \tag{2.1}$$

Proof. The argument here is an extension of the one given in Section 9 of [9]. Let

$\{\tilde{B}^x_t, x \in \mathbb{Z}^2\}$ be a system of (non-coalescing) independent rate one continuous time random walks with jump kernel p. Let $x_1, \dots, x_n \in \mathbb{Z}^2$ be distinct, set $x = (x_1, \dots, x_n)$ and define the non-collision event

$$D_t = \big\{\tilde{B}^{x_i}_s \neq \tilde{B}^{x_j}_s \text{ for all } s \le t \text{ and } i \neq j\big\}.$$

Clearly, if P_x is the law of $(\tilde{B}^{x_1}, \dots, \tilde{B}^{x_n})$, then $q_t(x) = P_x(D_t)$. Dependence on the fixed natural number n is suppressed.

By Lemma 9.12 of [9], there is a positive constant $C_{2.2}$ depending only on p and n such that for $x_1, \dots, x_n$ as above,

$$P_x(D_{2t}|D_t) = 1 - \frac{\binom{n}{2}\log 2}{\log t} + \frac{c(x,t)}{(\log t)^{3/2}}, \text{ where } |c(x,t)| \le C_{2.2} \qquad \text{whenever } \max_{i\neq j}\{|x_i - x_j|^4\} \vee e^4 \le t. \tag{2.2}$$

For $t \ge 1$ define $f_x(t) = (\log t)^{\binom{n}{2}} P_x(D_t)$ and $k(t) = \max\{i \ge 0 : 2^i \le t < 2^{i+1}\}$. Since $P_x(D_t)$ is decreasing in t, it is easy to see that

$$\frac{(k(t))^{\binom{n}{2}}}{(k(t)+1)^{\binom{n}{2}}} f_x(2^{k(t)+1}) \le f_x(t) \le \frac{(k(t)+1)^{\binom{n}{2}}}{(k(t))^{\binom{n}{2}}} f_x(2^{k(t)}). \tag{2.3}$$

For $m, m' \ge 1$, iterating conditional probabilities leads to

$$\begin{aligned} f_x(2^{m+m'}) &= f_x(2^m) \prod_{i=0}^{m'-1} \Big(1 + \frac{1}{m+i}\Big)^{\binom{n}{2}} P_x(D_{2^{m+i+1}}|D_{2^{m+i}}) \\ &= f_x(2^m) \prod_{k=m}^{m+m'-1} \Big(1 + \frac{1}{k}\Big)^{\binom{n}{2}} P_x(D_{2^{k+1}}|D_{2^k}). \end{aligned} \tag{2.4}$$

To make use of (2.4), we let $\bar{c}(x,k) = c(x, 2^k)/(\log 2)^{3/2}$, and note that it follows from (2.2) that

$$P_x(D_{2^{k+1}}|D_{2^k}) = 1 - \frac{\binom{n}{2}}{k} + \frac{\bar{c}(x,k)}{k^{3/2}}, \text{ where } |\bar{c}(x,k)| \le C_{2.5} \qquad \text{whenever } 2^k \ge \max_{i\neq j}\{|x_i - x_j|^4\} \vee e^4. \tag{2.5}$$

By the binomial theorem there is a constant $C_{2.6} > 0$ (depending only on n) so that

$$(1 + \frac{1}{k}\Big)^{\binom{n}{2}} = 1 + \frac{\binom{n}{2}}{k} + \frac{\underline{c}(k)}{k^2} \text{ where } \sup_{k\in\mathbb{N}} |\underline{c}(k)| \le C_{2.6}. \tag{2.6}$$

Taken together, the last two facts imply that for some $C_{2.7} > 0$, depending only on $p(\cdot)$ and n, there are constants $\tilde{c}(x,k)$ satisfying

$$\Big(1+\frac{1}{k}\Big)^{\binom{n}{2}} P_x(D_{2^{k+1}}|D_{2^k}) = 1+\frac{\tilde{c}(x,k)}{k^{3/2}} \text{ where } |\tilde{c}(x,k)| \le C_{2.7} \text{ for } 2^k \ge \max_{i\neq j}\{|x_i - x_j|^4\} \vee e^4. \tag{2.7}$$

Use this in (2.4) to see that, for all m and x satisfying $2^m \ge \max\{|x_i - x_j|^4\} \vee e^4$,

$$f_x(2^{m+m'}) = f_x(2^m) \prod_{k=m}^{m+m'-1} \Big(1 + \frac{\tilde{c}(x,k)}{k^{3/2}}\Big), \text{ where } \sup_{k\ge m} |\tilde{c}(x,k)| \le C_{2.7}. \tag{2.8}$$

If m also satisfies $m \geq j_0 \equiv \lceil (2C_{2.7})^{2/3} \rceil$, then the above bound implies

$$1 + \frac{\tilde{c}(x,k)}{k^{3/2}} \geq 1/2 \text{ if } k \geq m, \tag{2.9}$$

and in particular, each factor in the product in (2.8) is strictly positive.

Define $t_0 = e^4 \vee 2^{j_0}$ and $m_0(x) = \min\{m : 2^m \geq \max_{i \neq j}\{|x_i - x_j|^4 \vee t_0\} \geq j_0$. Then for any $m \geq m_0(x)$, (2.8) and (2.9) imply

$$\lim_{m' \to \infty} f_x(2^{m+m'}) = f_x(2^m) \prod_{k=m}^{\infty} \Big(1 + \frac{\tilde{c}(x,k)}{k^{3/2}}\Big) \tag{2.10}$$

exists, and is strictly positive. The fact that $\lim_{k\to\infty} f_x(2^k)$ exists combined with (2.3) actually proves (1.7), with

$$K_n(x_1, \dots, x_n) = f_x(2^m) \prod_{k=m}^{\infty} \Big(1 + \frac{\tilde{c}(x,k)}{k^{3/2}}\Big) \text{ for all } m \geq m_0(x). \tag{2.11}$$

It follows from this and (2.8) that for all $m \geq m_0(x) (\geq j_0)$,

$$\frac{f_x(2^m)}{K_n(x_1, \dots, x_n)} = \Big[\prod_{k=m}^{\infty} \Big(1 + \frac{\tilde{c}(x,k)}{k^{3/2}}\Big)\Big]^{-1} \leq \Big[\prod_{k=j_0}^{\infty} \Big(1 - \frac{C_{2.7}}{k^{3/2}}\Big)\Big]^{-1} < \infty, \tag{2.12}$$

where we have also used (2.9) and the definition of j_0.

Finally, if $t \geq 2(\max_{i \neq j} |x_i - x_j|^4 \vee t_0)$, then $2^{k(t)} \geq t/2$ implies $k(t) \geq m_0$, and by (2.12),

$$f_x(2^{k(t)}) \leq K_n(x_1, \dots, x_n) \Big[\prod_{k=j_0}^{\infty} \Big(1 - \frac{C_{2.7}}{k^{3/2}}\Big)\Big]^{-1}.$$

It now follows from (2.3) and (2.12) that for such t (note also $k(t) \geq 1$),

$$f_x(t) \leq \frac{(k(t)+1)^{\binom{n}{2}}}{(k(t))^{\binom{n}{2}}} f_x(2^{k(t)}) \leq 2^{\binom{n}{2}} \Big[\prod_{k=j_0}^{\infty} \Big(1 - \frac{C_{2.6}}{k^{3/2}}\Big)\Big]^{-1} K_n(x_1, \dots, x_n). \tag{2.13}$$

This completes the proof of Lemma 2.1. $\square$

For disjoint finite nonempty sets of $\mathbb{Z}^2$, $A_1, \dots A_n$, introduce

$$\Gamma_t = \big\{\sigma(A_1, \dots, A_n) > t, \tau(A_1, \dots, A_n) < t\big\},$$

and define $q_{A_1,\dots,A_n}(t) = \widehat{P}(\Gamma_t)$. Note that if $A_i = \{a_i\}$ are all singletons, then $\tau(A_1, \dots, A_n) = 0$, and so $q_{A_1,\dots,A_n}(t) = q_{a_1,\dots,a_n}(t)$ agrees with our earlier notation.

Proof of Proposition 1.6. Assume A_i, $i \leq n$ and $a_i \in A_i$ are as in Proposition 1.6, and define $K_n(A_1, \dots, A_n)(\leq \infty)$ as in (1.9). On the event Γ_t, $B_t^{a_i} \neq B_t^{a_j}$ for all $i \neq j$, which implies $q_{A_1,\dots,A_n}(t) \leq q_{a_1,\dots,a_n}(t)$. It follows from (1.7) that

$$\limsup_{t\to\infty} (\log t)^{\binom{n}{2}} q_{A_1,\dots,A_n}(t) \leq K_n(a_1, \dots, a_n) < \infty. \tag{2.14}$$

Letting $\tau^* = \tau(A_1, \dots, A_n)$ and $\sigma^* = \sigma(A_1, \dots, A_n)$, and applying the Markov property at time τ^*, we have

$$\begin{aligned}&(\log t)^{\binom{n}{2}} q_{A_1,\dots,A_n}(t)\\ &= \sum_{\text{distinct } x_1,\dots,x_n \in \mathbb{Z}^2} \int_0^t \widehat{P}(\sigma^* > \tau^* \in du, B_{\tau^*}^{a_i} = x_i \text{ for } 1 \leq i \leq n)(\log t)^{\binom{n}{2}} q_{x_1,\dots,x_n}(t-u).\end{aligned} \tag{2.15}$$

By (1.7), for fixed $u \in [0,t)$, and distinct $x_1, \dots, x_n \in \mathbb{Z}^2$,

$$\lim_{t\to\infty} (\log t)^{\binom{n}{2}} q_{x_1,\dots,x_n}(t-u) = K_n(x_1,\dots,x_n). \tag{2.16}$$

In view of the definition of $K_n(A_1,\dots,A_n)$, (2.14), (2.15), and (2.16), Fatou's Lemma implies that

$$\begin{aligned} K_n(A_1,\dots,A_n) \le \liminf_{t\to\infty} (\log t)^{\binom{n}{2}} q_{A_1,\dots,A_n}(t) &\le \limsup_{t\to\infty} (\log t)^{\binom{n}{2}} q_{A_1,\dots,A_n}(t) \\ &\le K_n(a_1,\dots,a_n) < \infty. \end{aligned} \tag{2.17}$$

Now introduce

$$\Delta^* = \max_{i\neq j}\{|B^{a_i}_{\tau^*} - B^{a_j}_{\tau^*}|^4\},$$

and the disjoint decomposition $\Gamma_t = \Gamma_t^{(1)} \cup \Gamma_t^{(2)} \cup \Gamma_t^{(3)}$, where

$$\begin{aligned} \Gamma_t^{(1)} &= \{\sigma^* > t, \tau^* \in (t^{1/3}, t]\} \\ \Gamma_t^{(2)} &= \{\sigma^* > t, \tau^* \le t^{1/3}, \Delta^* > t/4\} \\ \Gamma_t^{(3)} &= \{\sigma^* > t, \tau^* \le t^{1/3}, \Delta^* \le t/4\}. \end{aligned}$$

We will show that as $t \to \infty$,

$$(\log t)^{\binom{n}{2}} \widehat{P}(\Gamma_t^{(i)}) \to 0 \text{ for } i = 1,2, \text{ and} \tag{2.18}$$

$$(\log t)^{\binom{n}{2}} \widehat{P}(\Gamma_t^{(3)}) \to K_n(A_1,\dots,A_n), \tag{2.19}$$

proving (1.8). It is in the proof of (2.19) that we use Lemma 2.1.

On the event $\Gamma_t^{(1)}$, $\sum_{i=1}^n |B^{A_i}_{t^{1/3}}| \ge n+1$ (recall that $|A_1| + \cdots + |A_n| \ge n+1$). It follows that if $A = \cup_{i=1}^n A_i$, then

$$(\log t)^{\binom{n}{2}} \widehat{P}(\Gamma_t^{(1)}) \le \sum_{\text{distinct } y_0,y_1,\dots,y_n \in A} (\log t)^{\binom{n}{2}} q_{y_0,y_1,\dots,y_n}(t^{1/3}) \to 0 \text{ as } t \to \infty, \tag{2.20}$$

since each probability in the sum is $O\big((\log t)^{-\binom{n+1}{2}}\big)$ by (1.7).

On the event $\Gamma_t^{(2)}$, there must exist $a \neq b \in A$ such that $|B^a_u - B^b_u| > (t/4)^{1/4}$ for some $u \le t^{1/3}$. Assuming $(t/4)^{1/4} > 2 \max_{a\neq b\in A} |a-b|$, we have the crude bound

$$\begin{aligned} \widehat{P}(\Gamma_t^{(2)}) &\le \sum_{a\neq b\in A} \widehat{P}\Big(\sup_{0\le u\le t^{1/3}} |B^a_u - B^b_u| \ge (\tfrac{t}{4})^{1/4}\Big) \\ &= \sum_{a\neq b\in A} \widehat{P}\Big(\sup_{0\le u\le t^{1/3}} |B^0_{2u}| \ge \tfrac12 (\tfrac{t}{4})^{1/4}\Big) \\ &\le \binom{|A|}{2} \widehat{P}\Big(\sup_{0\le u\le 2t^{1/3}} |B^0_u|^2 \ge \tfrac14 (\tfrac{t}{4})^{1/2}\Big). \end{aligned}$$

Doob's inequality implies

$$(\log t)^{\binom{n}{2}} \widehat{P}(\Gamma_t^{(2)}) \le \binom{|A|}{2} (\log t)^{\binom{n}{2}} \frac{\widehat{E}(|B^0_{2t^{1/3}}|^2)}{\frac14 (t/4)^{1/2}} = 32\sigma^2 \binom{|A|}{2} (\log t)^{\binom{n}{2}} t^{-1/6} \to 0 \text{ as } t \to \infty,$$

completing the proof of (2.18)

To handle $\Gamma_t^{(3)}$, we recall (2.1) and suppose that $t > 4t_0 \vee 8$ (so that $t - t^{1/3} > t/2$). Applying the Markov property at time τ^* gives us

$$(\log t)^{\binom{n}{2}} \widehat{P}(\Gamma_t^{(3)})$$
$$= \sum_{\substack{\text{distinct } x_1,\dots,x_n \in \mathbb{Z}^2 \\ \max_{i\neq j}\{|x_i - x_j|^4\} \le t/4}} \int_0^{t^{1/3}} \widehat{P}(\sigma^* > \tau^* \in du, B^{a_i}_{\tau^*} = x_i \text{ for } 1 \le i \le n)(\log t)^{\binom{n}{2}} q_{x_1,\dots,x_n}(t-u). \tag{2.21}$$

For $u \le t^{1/3}$ and $t/4 \ge \max_{i\neq j}\{|x_i - x_j|^4\}$ we have (use also $t > 4t_0 \vee 8$)

$$(t-u)/2 > (t - t^{1/3})/2 > t/4 \ge t_0 \vee \max_{i \neq j}\{|x_i - x_j|^4\},$$

and so Lemma 2.1 applies to show that

$$(\log t)^{\binom{n}{2}} q_{x_1,\dots,x_n}(t-u) \le \frac{(\log t)^{\binom{n}{2}}}{(\log(t-u))^{\binom{n}{2}}} C_{2.1} K_n(x_1,\dots,x_n) \le 2^{\binom{n}{2}} C_{2.1} K_n(x_1,\dots,x_n).$$

For the last inequality use $t - u > t/2$ and $t > 8$. The right-hand side is integrable with respect to $\widehat{P}(\sigma^* > \tau^* \in du)1(x_1,\dots,x_n \text{ distinct})\, d\lambda$ for counting measure λ on $(\mathbb{Z}^2)^n$, by (2.17) and the definition of $K_n(A_1,\dots,A_n)$. So we can use (2.16) and dominated convergence to take the limit as $t \to \infty$ inside the integral in (2.21) and hence, recalling again the definition of $K_n(A_1,\dots,A_n)$, prove (2.19) and so complete the proof. □

3 Cancellative processes, voter model perturbations and examples

3.1 Cancellative processes

Let $\mathcal{N}$ be a non-empty finite subset of $\mathbb{Z}^d \setminus \{0\}$, and call such a subset a *general* neighbourhood. Our cancellative processes are translation invariant spin systems with rate functions satisfying

$$c(x,\xi) = \frac{k_0}{2}\left[1 - (2\xi(x) - 1)\sum_{A \subset \bar{\mathcal{N}}} \beta_0(A) H(\xi, A + x)\right], \tag{3.1}$$

where $H(\xi, A) = \prod_{y \in A}(2\xi(y) - 1)$, k_0 is a positive constant, $\beta_0 \ge 0$, $\beta_0(\varnothing) = 0$, and $\sum_{A \subset \bar{\mathcal{N}}} \beta_0(A) = 1$. This last implies **1** is a trap for ξ_t, that is, $c(x, \mathbf{1}) \equiv 0$. The restriction to subsets A of $\bar{\mathcal{N}}$ means our definition is a bit more restrictive than that in (1.16) of [13] or (4.4) in Section III.4 of [24]. A recent summary of properties of cancellative processes is given in Sections 1 and 2 of [13]. We say ξ is a *good* cancellative process if, in addition, $\beta_0(A) > 0$ for some A with $|A| > 1$.

Remark 3.1. For any probability q_0 on $\mathcal{N}$, the voter model with kernel q_0 is cancellative with $k_0 = 1$, and $\beta_0(A) = q_0(y)$ if $A = \{y\}$, and zero if $|A| \ne 1$, as one can easily check (or see Example III.4.16 in [24]). So the voter model is cancellative, but not a good cancellative process. The converse also holds and equally easy to check: If ξ is cancellative but not good cancellative, then ξ has rates equal to k_0 times those of a voter model with kernel $q_0(y) = \beta_0(\{y\})$ $(y \in \mathcal{N})$.

Another useful condition, which will hold for our main results, is

$$\beta_0(A) = 0 \text{ if } |A| \text{ is even}, \tag{3.2}$$

which by Lemma 2.1 in [13] is equivalent to

$$\mathbf{0} \text{ is a trap for } \xi_t, \tag{3.3}$$

and also to

$$\xi \text{ is } 0-1 \text{ symmetric, that is, } c(x,\xi)=c(x,\widehat{\xi}). \tag{3.4}$$

A cancellative process has a translation invariant stationary distribution, which is the weak limit of the process started in Bernoulli product measure with density $1/2$ (see Corolllary III.1.8 of [20]). Under the above symmetry it will have density $\mathbf{1}/2$ and so we denote it by $\nu_{1/2}$. A simple proof of the existence of $\nu_{1/2}$ under (3.4) is given after Lemma 2.1 in [13]. It is possible that $\nu_{1/2}=(1/2)(\delta_{\mathbf{0}}+\delta_{\mathbf{1}})$, and so it need not have the coexistence property. For example, this is the case for the voter model when $d\le 2$ (Corollary V.1.13 of [24]).

As defined in [5], a general neighbourhood $\mathcal{N}$ and nonnegative sequence $a=(a_\ell)$, $1\le \ell\le |\mathcal{N}|$, $(a_l)\not\equiv 0$, defines a nonlinear voter model $\xi_t:\mathbb{Z}^d\to\{0,1\}$ if the rate at which the opinion at a site flips to the opposite opinion is a_ℓ if ℓ of its neighbours hold this opposite opinion. That is, ξ_t is the spin-flip system defined by the rate function

$$c(x,\xi)=\widehat{\xi}(x)\sum_{\ell=1}^{|\mathcal{N}|}a_\ell 1\{n_1(x,\xi)=\ell\}+\xi(x)\sum_{\ell=1}^{|\mathcal{N}|}a_\ell 1\{n_0(x,\xi)=\ell\},$$

where $n_i(x,\xi)=\sum_{y\in x+\mathcal{N}}1\{\xi(y)=i\}$, $i=1,2$. Clearly every nonlinear voter model rate function $c(x,\xi)$ satisfies the symmetry condition (3.4). In [5] $\mathcal{N}$ satisfied additional symmetry and irreducibility conditions, but they are not needed for the results in this section. Henceforth we exclude the trivial case of $a\equiv 0$ from consideration. The q-voter model on $\mathbb{Z}^d$ (with neighbourhood $\mathcal{N}$) is the nonlinear voter model with $a_\ell=(\ell/|\mathcal{N}|)^q$.

The following result, from page 129 of [5] provides a means of checking that a nonlinear voter model is cancellative without checking (3.1) directly. We give the short proof for completeness and to highlight the choice of β_0 and k_0 which will enter later.

Proposition 3.2. *Let $\mathcal{N}$ be a general neighbourhood and $a_\ell\ge 0$ for $\ell=1,\dots,|\mathcal{N}|$ with $(a_\ell)\not\equiv 0$. Define the $|\mathcal{N}|\times|\mathcal{N}|$ matrix $\mathbf{M}$ by*

$$M(k,j)=\sum_{odd\ i\le j\wedge k}\binom{j}{i}\binom{|\mathcal{N}|-j}{k-i},\qquad 1\le k,j\le|\mathcal{N}|. \tag{3.5}$$

If there is a nonnegative sequence $\alpha=(\alpha_k)$, $1\le k\le|\mathcal{N}|$, such that

$$a_\ell=\sum_{k=1}^{|\mathcal{N}|}\alpha_k\,M(k,\ell),\quad 1\le\ell\le|\mathcal{N}|, \tag{3.6}$$

then the nonlinear voter model determined by (a_ℓ) is a ***cancellative*** *nonlinear voter model. Moreover in (3.1) $k_0=\sum_{\ell=1}^{|\mathcal{N}|}\alpha_\ell$ and for $A\subset\bar{\mathcal{N}}$,*

$$\beta_0(A)=\frac{1}{k_0}\times\begin{cases}0 & \textit{if } A=\{0\} \textit{ or } |A| \textit{ is even}\\ \alpha_m & \textit{if } 0\notin A, |A|=m \textit{ is odd}\\ \alpha_{m-1} & \textit{if } 0\in A, |A|=m \textit{ is odd,}\end{cases} \tag{3.7}$$

and so (3.2) also holds.

Proof. Define $\xi(A)=\sum_{a\in A}\xi(a)$ and $\widehat{\xi}(A)=\sum_{a\in A}\widehat{\xi}(a)$. In (3.1) since $2\xi(x)-1=(-1)^{\widehat{\xi}(x)}$ and $\sum_A\beta_0(A)=1$, we can rewrite (3.1) for $x=0$ as

$$c(0,\xi)=\frac{k_0}{2}\sum_{A\subset\bar{\mathcal{N}}}\beta_0(A)\Big[1-(-1)^{\widehat{\xi}(A)+\widehat{\xi}(0)}\Big]. \tag{3.8}$$

Let $\boldsymbol{\alpha} = (\alpha_j)$ be a given nonnegative sequence such that (3.6) defines non-negative a_k, not all identically zero. Let $k_0 = \sum_{\ell=1}^{|\mathcal{N}|} \alpha_\ell > 0$, the latter since otherwise $a \equiv 0$. For $A \subset \bar{\mathcal{N}}$, define $\beta_0(A)$ by (3.7). Then $\beta_0(\emptyset) = 0$ and $\sum_{A\subset\bar{\mathcal{N}}} \beta_0(A) = \frac{1}{k_0}\sum_{\ell=1}^{|\mathcal{N}|} \alpha_\ell = 1$. Note also that (3.2) holds and hence so does (3.4). With these choices, (3.8) becomes

$$c(0,\xi) = \frac{1}{2}\sum_{m \text{ odd}} \Big\{\alpha_m \sum_{A\not\ni 0, |A|=m} + \mathbf{1}(m\ge 3)\alpha_{m-1}\sum_{A\ni 0,|A|=m}\Big\}\Big[1-(-1)^{\widehat{\xi}(A)+\widehat{\xi}(0)}\Big]. \tag{3.9}$$

Observe that in the second sum, $m \ge 3$ because we have set $\beta(\{0\}) = 0$.

Observe that if m is odd and $|A| = m$, if $\xi(0) = 0$, then

$$\frac{1}{2}\Big[1-(-1)^{\widehat{\xi}(A)+\widehat{\xi}(0)}\Big] = \frac{1}{2}\Big[1+(-1)^{\widehat{\xi}(A)}\Big] = 1\{\widehat{\xi}(A) \text{ is even}\} = 1\{\xi(A) \text{ is odd}\},$$

where we have used the fact that $\xi(A) + \widehat{\xi}(A) = |A|$. Therefore,

$$c(0,\xi) = \sum_{m \text{ odd}} \Big\{\alpha_m \sum_{A\not\ni 0, |A|=m} + \mathbf{1}(m\ge 3)\alpha_{m-1}\sum_{A\ni 0,|A|=m}\Big\} 1\{\xi(A) \text{ is odd}\}. \tag{3.10}$$

For any ξ, odd m, $1 \le j \le |\mathcal{N}|$, define

$$\mathcal{A}'_m(j) = \Big\{A \subset \bar{\mathcal{N}} : 0 \notin A, |A| = m, \xi(\mathcal{N}) = j, \xi(A) \text{ is odd}\Big\}$$
$$\mathcal{A}''_m(j) = \Big\{A \subset \bar{\mathcal{N}} : 0 \in A, |A| = m, \xi(\mathcal{N}) = j, \xi(A) \text{ is odd}\Big\}.$$

Then

$$c(0,\xi) = \sum_{m \text{ odd}} \alpha_m |\mathcal{A}'_m(j)| + \sum_{m\ge 3, \text{ odd}} \alpha_{m-1}|\mathcal{A}''_m(j)| \qquad \text{if } \xi(0) = 0,\ \xi(\mathcal{N}) = j. \tag{3.11}$$

Continue to assume $\xi(\mathcal{N}) (= n_1(0,\xi)) = j$ and $\xi(0) = 0$, consider $A \in \mathcal{A}'_m(j)$, and the disjoint union

$$A = (A \cap \{x : \xi(x) = 1\}) \cup (A \cap \{x : \xi(x) = 0\}).$$

By counting the number of ways to choose the first set with the requirement $i = \xi(A) \le j \wedge m$ odd, we find that

$$|\mathcal{A}'_m(j)| = \sum_{\substack{i \text{ odd}\\ i\le j\wedge m}} \binom{j}{i}\binom{|\mathcal{N}|-j}{m-i}.$$

Here we build $A \subset \mathcal{N}$ by first choosing the i sites of A to put 1's from the available j sites in state 1, and then choose the $m-i$ sites of A to put 0's from the $|\mathcal{N}| - j$ available sites in state 0. Similar reasoning (recall $\xi(0) = 0$ and now 0 must be in A) leads to

$$|\mathcal{A}''_m(j)| = \sum_{\substack{i \text{ odd}\\ i\le j\wedge(m-1)}} \binom{j}{i}\binom{|\mathcal{N}|-j}{m-1-i}.$$

Insert the above into (3.11) to get for $\xi(0) = 0,\ \xi(\mathcal{N}) = j$,

$$\begin{aligned} c(0,\xi) &= \sum_{1\le k\le|\mathcal{N}|} \alpha_k \sum_{\substack{i \text{ odd}\\ i\le j\wedge k}} \binom{j}{i}\binom{|\mathcal{N}|-j}{k-i} \\ &= \sum_{k=1}^{|\mathcal{N}|} \alpha_k M(k,j) = a_j. \end{aligned}$$

If $\xi(0) = 1$, and $\widehat{\xi}(\mathcal{N})(= n_0(0,\xi)) = j$, then $\widehat{\xi}(0) = 0$. Use the symmetry (3.4) noted earlier, and the above with $\widehat{\xi}$ in place of ξ to get

$$c(0,\xi) = c(0,\widehat{\xi}) = a_j.$$

Therefore the cancellative system corresponding to the above β_0 and k_0 is indeed the non-linear voter model determined by (a_j). □

Lemma 3.3. *Assume $d \geq 1$, and $\mathcal{N}$ is a general neighbourhood with $2 \leq |\mathcal{N}| \leq 8$.*
(a) For $2 \leq |\mathcal{N}| \leq 4$ and all $q \in [0,1]$, the corresponding q-voter model on $\mathbb{Z}^d$ is cancellative, and for $q \in [0,1)$, in (3.1) *we have $\beta_0(A) > 0$ for all $A \subset \bar{\mathcal{N}}$ with $|A| = 3$.*
(b) For q sufficiently close to 1, the corresponding q-voter model on $\mathbb{Z}^d$ is cancellative for $q \leq 1$ and the last conclusion in (a) holds for $q < 1$.

Proof. (a) Suppose first that $|\mathcal{N}| = 4$. it is straightforward to check that

$$\mathbf{M} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 3 & 0 \\ 3 & 2 & 1 & 4 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{M}^{-1} = \frac{1}{8}\begin{bmatrix} -2 & 0 & 2 & 4 \\ 0 & 2 & 0 & -6 \\ 2 & 0 & -2 & 4 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$

Thus, by (3.6), for $a_\ell = (\ell/|\mathcal{N}|)^q$, $\boldsymbol{\alpha} = \mathbf{a}\mathbf{M}^{-1}$ is given by

$$\begin{aligned} \alpha_1(q) &= -\frac{1}{4}\Big(\frac{1}{4}\Big)^q + \frac{1}{4}\Big(\frac{3}{4}\Big)^q + \frac{1}{8} \\ \alpha_2(q) &= \frac{1}{4}\Big(\frac{1}{2}\Big)^q - \frac{1}{8} \\ \alpha_3(q) &= \frac{1}{4}\Big(\frac{1}{4}\Big)^q - \frac{1}{4}\Big(\frac{3}{4}\Big)^q + \frac{1}{8} \\ \alpha_4(q) &= \frac{1}{2}\Big(\frac{1}{4}\Big)^q - \frac{3}{4}\Big(\frac{1}{2}\Big)^q + \frac{1}{2}\Big(\frac{3}{4}\Big)^q - \frac{1}{8}. \end{aligned}$$

It is now an enjoyable calculus exercise to check that each $\alpha_\ell(q) \geq 0$ for all $0 \leq q \leq 1$ and each $\alpha_\ell(q) > 0$ for $0 \leq q < 1$. By Proposition 3.2 the cancellative property holds for $q \in [0,1]$, and (3.7), together with $\alpha_2, \alpha_3 > 0$, give the final conclusion for any $q \in [0,1)$. The cases $|\mathcal{N}| = 2,3$ are handled the same way, with simpler calculations. This prove (a).

(b) Consider $|\mathcal{N}| = 8$. It is still easy to write down $\mathbf{M}$ from (3.5). Then `maple` can be used to find $\mathbf{M}^{-1}$ (see the Appendix), and (3.6) can be used to write down each $\alpha_\ell(q)$, $1 \leq \ell \leq |\mathcal{N}|$ explicitly. With these formulas it is easy to check that $\alpha_1(q) \geq 2^{-|\mathcal{N}|+1}$ for $0 \leq q \leq 1$, and also that $\alpha_\ell(1) = 0$ for $2 \leq \ell \leq |\mathcal{N}|$. Some simple, if lengthy, arithmetic shows that

$$\alpha_\ell'(1) = -2^{-k_\ell}\log(\frac{n_\ell}{m_\ell}) < 0 \text{ for } 2 \leq \ell \leq |\mathcal{N}|, \tag{3.12}$$

where k_ℓ, m_ℓ, n_ℓ are positive integers with $m_\ell < n_\ell$ and $k_\ell \leq |\mathcal{N}| + 1$. All of these quantities are given in the Appendix below. It follows that each $\alpha_\ell(q)$ must be strictly positive for $q < 1$ sufficiently close to 1. Thus Proposition 3.2 implies the cancellative property for $q \leq 1$ close to 1, and (3.7) implies that $\beta_0(A) > 0$ if $|A| = 3$, $A \subset \bar{\mathcal{N}}$ for $q < 1$ close to 1.

The cases $|\mathcal{N}| = 5,6,7$ are handled the same way. We omit the details. □

Conjecture 3.4. *All q-voter models are cancellative for any $d \geq 1$, any neighbourhood $\mathcal{N}$, $|\mathcal{N}| \geq 2$, and any $0 \leq q \leq 1$.*

3.2 Finite range voter model perturbations and examples

For $d \geq 2$, let the probability kernel $p : \mathbb{Z}^d \to [0,1]$, the local frequencies of type i, $f_i(x,\xi)$, and the voter model rates $c^{\mathsf{vm}}(x,\xi)$ be as in Section 1.4. In particular,

$$p \text{ has covariance matrix } \sigma^2 I \text{ for some } \sigma > 0. \tag{3.13}$$

Consider a neighbourhood $\mathcal{N}$ containing the support of p and recall the definition of a voter model perturbation with finite range in $\mathcal{N}$, given in Definition 1.8. Recall also from (1.18) the asymptotic rate function $r_s(A) = g_1(1_A)$ for $A \subset \mathcal{N}$.

It follows from (1.14) or (1.15) that by decreasing ε_0, if necessary, we may assume that

$$\sup_{0<\varepsilon\leq\varepsilon_0} \|g_0^\varepsilon\|_\infty + \|g_1^\varepsilon\|_\infty = C(g) < \infty. \tag{3.14}$$

Remark 3.5. For $d \geq 3$ Proposition 1.1 of [7] shows that in the finite range case (both voter model and perturbation depend only on sites in $\mathcal{N}$), our definition of finite range voter model perturbation coincides with the voter perturbations considered in [13] for $d \geq 3$ (i.e. those satisfying (1.10)-(1.15) of that reference). The notation here is a bit different as what we call g_i^ε is denoted by h_i^ε in [7], where a related quantity is called g_i^ε and is non-negative. This non-negativity plays a role in the definition of the dual process used to study $\xi^{[\varepsilon]}$. We will only use the dual process implicitly when quoting arguments from [13] in Section 4 so there should be no confusion.

For $d = 2$ to obtain a scaling limit theorem to super-Brownian motion we have added the condition that $\mathbf{1}$ is a trap. That there is distinct behavior if $\mathbf{1}$ is not a trap is demonstrated by Theorem 1.3 of [17] where for the 2-dimensional contact process with rapid voting (a voter model perturbation) it is shown that the critical birth rate for survival must diverge to $+\infty$ as the voter rate gets large. We also have dropped the Hölder convergence rate in (1.14) which entered in the pde analysis used in earlier work but will not be needed here. Indeed, the Hölder convergence rate will not be satisfied by the 2-dimensional Lotka-Volterra models described below.

In two dimensions, recall from Definition 1.14 the notion of a finite range voter model perturbation which is asymptotically symmetric, and which will play an important role in our scaling limit theorems.

Remark 3.6. Clearly $0-1$ symmetry of the voter model perturbations c_ε (for all ε) implies asymptotic symmetry with $r^a = g^a \equiv 0$. Note also that by (1.14) (or (1.15)) and (1.28), asymptotic symmetry of a voter model perturbation implies

$$g_1(\xi) = g_0(\widehat{\xi}). \tag{3.15}$$

For $d \geq 3$ finite range voter model perturbations rescale to SBM if time is rescaled by $1/\varepsilon$ and space was scaled down by $\sqrt{\varepsilon}$ (see [10]). If $d = 2$ these scaling parameters will change, as already noted in [9] in the special case of Lotka-Volterra models. Here the scaling parameter $N > e^3$ is the unique solution $N = N(\varepsilon)$ of

$$\varepsilon = \frac{(\log N)^3}{N},\ \varepsilon \in (0, \varepsilon_0]. \tag{3.16}$$

It is a calculus exercise to verify the existence and uniqueness of such an N for $\varepsilon \in (0,1] \supset (0,\varepsilon_0]$. Henceforth for $d = 2$ we usually consider N as our fundamental parameter and set $\varepsilon = \varepsilon_N = (\log N)^3/N$. The constraint $\varepsilon \leq \varepsilon_0$ leads to

$$N \geq N(\varepsilon_0) > e^3, \tag{3.17}$$

where the actual value of $N(\varepsilon_0)$ is of little concern as we will be interested in taking $N \to \infty$.

Proceeding now with $d=2$, our rescaled lattice is $S_N=\mathbb{Z}^2/\sqrt{N}$. For $\xi\in\{0,1\}^{\mathbb{Z}^2}$, define the rescaled state $\xi^{(N)}\in\{0,1\}^{S_N}$ by $\xi^{(N)}(x)=\xi(x\sqrt{N})$. Let $\{c_\varepsilon,\ \varepsilon\le\varepsilon_0\}$, be an asymptotically symmetric finite range voter model perturbation and consider the rescaled rate function

$$c^N(x,\xi^{(N)})=Nc_{\varepsilon_N}(x\sqrt{N},\xi),\ x\in S_N,\ \xi\in\{0,1\}^{\mathbb{Z}^2} \tag{3.18}$$

for the rescaled voter model perturbation process

$$\xi_t^N(x)=\xi_{tN}^{[\varepsilon_N]}(x\sqrt{N}),\ x\in S_N. \tag{3.19}$$

For $x\in S_N$, and $\xi\in\{0,1\}^{\mathbb{Z}^2}$ introduce

$$c^{N,\mathsf{vm}}(x,\xi^{(N)})=c^{\mathsf{vm}}(x\sqrt{N},\xi),$$

$$c^{N,s}(x,\xi^{(N)})=\widehat{\xi}(x\sqrt{N})g_0^{\varepsilon_N}(\hat{\xi}|_{x\sqrt{N}+\mathcal{N}})+\xi(x\sqrt{N})g_0^{\varepsilon_N}(\xi|_{x\sqrt{N}+\mathcal{N}}), \tag{3.20}$$

and

$$c^{N,a}(x,\xi^{(N)})=\widehat{\xi}(x\sqrt{N})(\log N)^2[g_1^{\varepsilon_N}(\xi|_{x\sqrt{N}+\mathcal{N}})-g_0^{\varepsilon_N}(\widehat{\xi}|_{x\sqrt{N}+\mathcal{N}})]. \tag{3.21}$$

Then for $x\in S_N$ and $\xi\in\{0,1\}^{\mathbb{Z}^2}$,

$$c^N(x,\xi^{(N)})=Nc^{\mathsf{vm}}(x\sqrt{N},\xi)+(\log N)^3c^*_{\varepsilon_N}(x\sqrt{N},\xi) \tag{3.22}$$

$$=Nc^{N,\mathsf{vm}}(x,\xi^{(N)})+(\log N)c^{N,a}(x,\xi^{(N)})+(\log N)^3c^{N,s}(x,\xi^{(N)}). \tag{3.23}$$

Clearly $c^{N,s}$ is symmetric, that is

$$c^{N,s}(x,\xi^{(N)})=c^{N,s}(x,\hat{\xi}^{(N)}), \tag{3.24}$$

and a short calculation shows that

$$\text{if } c^*_\varepsilon(x,\xi)=c^*_\varepsilon(x,\widehat{\xi}) \text{ for all } \varepsilon,x,\xi, \text{ then } c^{N,a}\equiv 0 \text{ for all } N. \tag{3.25}$$

Moreover it is easy to check that the decomposition in (3.23) uniquely determines $c^{N,s}$ and $c^{N,a}$ if we assume that $c^{N,s}$ is symmetric and for every $x\in S_N$, $c^{N,a}(x,\xi^{(N)})=0$ if $\xi^{(N)}(x)=1$. If for $x\in\mathbb{Z}^2$ and $\xi\in\{0,1\}^{\mathbb{Z}^2}$,

$$c^s(x,\xi)=\widehat{\xi}(x)g_0(\widehat{\xi}|_{x+\mathcal{N}})+\xi(x)g_0(\xi|_{x+\mathcal{N}}),\ \text{and } c^a(x,\xi)=\widehat{\xi}(x)g^a(\xi|_{x+\mathcal{N}}),$$

then c^s is symmetric, and from (3.20), (3.21), (1.15), $\log(1/\varepsilon_N)/\log N\to 1$, and (1.28) we have

$$\sup_{x\in\mathbb{Z}^2,\xi\in\{0,1\}^{\mathbb{Z}^2}}|c^{N,s}(x/\sqrt{N},\xi^{(N)})-c^s(x,\xi)|+|c^{N,a}(x/\sqrt{N},\xi^{(N)})-c^a(x,\xi)|\to 0 \text{ as } N\to\infty. \tag{3.26}$$

Use the definition of $c^{N,a}$ and the convergence in (1.28) to see that if

$$r^{N,a}(A)=(\log N)^2[g_1^{\varepsilon_N}(1_A)-g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A})] \text{ for } A\subset\mathcal{N},$$

then

$$c^{N,a}(x/\sqrt{N},\xi^{(N)})=\widehat{\xi}(x)\sum_{\emptyset\neq A\subset\mathcal{N}}r^{N,a}(A)1\{\xi|_{x+\mathcal{N}}=1_{x+A}\} \tag{3.27}$$

and

$$\lim_{N\to\infty}r^{N,a}(A)=g^a(1_A):=r^a(A) \text{ for all } A\subset\mathcal{N}. \tag{3.28}$$

The fact that $\mathbf{0}$ and $\mathbf{1}$ are traps implies $g_1^{\varepsilon_N}(1_\emptyset) = g_0^{\varepsilon_N}(1_{\mathcal{N}}) = 0$ and so we have dropped the $A = \emptyset$ term in the sum. Similarly if

$$r^{N,s}(A) = g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A}) \text{ for } A \subset \mathcal{N},$$

then we have

$$c^{N,s}(x/\sqrt{N}, \xi^{(N)}) = \widehat{\xi}(x) \sum_{\emptyset \neq A \subset \mathcal{N}} r^{N,s}(A) 1\{\xi|_{x+\mathcal{N}} = 1_{x+A}\} + \xi(x) \sum_{\emptyset \neq A \subset \mathcal{N}} r^{N,s}(A) 1\{\xi|_{x+\mathcal{N}} = 1_{x+\mathcal{N}\setminus A}\} \tag{3.29}$$

and (recall also (3.15))

$$\lim_{N\to\infty} r^{N,s}(A) = g_0(1_{\mathcal{N}\setminus A}) = g_1(1_A) := r^s(A) \text{ for all } A \subset \mathcal{N}. \tag{3.30}$$

It follows from (3.14), (3.28) and (3.30) that

$$\|r\| := \sup_{N \geq N(\varepsilon_0),\ \emptyset \neq A \subset \mathcal{N}} |r^{N,a}(A)| \vee |r^{N,s}(A)| < \infty. \tag{3.31}$$

The following result will be useful for some comparison results in Section 7.1.

Lemma 3.7. *For all non-empty $A \subset \mathcal{N}$, $p(A) = 0$ implies that $r^{N,s}(A) \geq 0$ for all N.*

Proof. Assume $p(A) = 0$ for a fixed set A as above. We must show that

$$g_0^\varepsilon(1_{\mathcal{N}\setminus A}) \geq 0 \quad \text{for all } \varepsilon \in [0, \varepsilon_0]. \tag{3.32}$$

Choose $\xi \in \{0,1\}^{\mathbb{Z}^2}$ such that $\xi(0) = 1$ and $\xi|_{\mathcal{N}} = 1_{\mathcal{N}\setminus A}$. Then

$$0 \leq c_\varepsilon(0,\xi) = f_0(0,\xi) + g_0^\varepsilon(1_{\mathcal{N}\setminus A}) = \sum_{y\in A} p(y)\widehat{\xi}(y) + g_0^\varepsilon(1_{\mathcal{N}\setminus A}) = g_0^\varepsilon(1_{\mathcal{N}\setminus A}).$$

The next to last equality holds because $\widehat{\xi}(y) = 0$ for $y \notin A$, and the last equality holds by our assumption on A. This proves (3.32). ☐

Recall the notation $\Theta^\pm(A)$ from (1.10) and K_2 from Proposition 1.6. We also recall Θ_3 from (1.19) and Θ_2 from (1.31):

$$\Theta_3 = \sum_{\emptyset \neq A \subset \mathcal{N}} r^s(A)(\Theta^+(A) - \Theta^-(A)) \quad \text{and} \quad \Theta_2 = \sum_{\emptyset \neq A \subset \mathcal{N}} r^a(A) K_2(A, \bar{\mathcal{N}} \setminus A). \tag{3.33}$$

The following identities (see Remark 5.4 below) will simplify Θ_3 in some of the examples we now discuss:

$$\sum_{\emptyset \neq A \subset \mathcal{N}} |A|(\Theta^+(A) - \Theta^-(A)) = 0, \tag{3.34}$$

and

$$\sum_{\emptyset \neq A \subset \mathcal{N}} (\Theta^+(A) - \Theta^-(A)) = \kappa = \lim_{t\to\infty} (\log t)^3 \hat{P}(|B_t^{\bar{\mathcal{N}}}| = 3) > 0. \tag{3.35}$$

Example 3.8. (q-Voter Models) We begin with all $d \geq 2$, follow [4] and consider, as in the Introduction, the q-voter models with kernel $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, for some neighbourhood $\mathcal{N}$. We are interested in q near 1, $q < 1$, and so let $q = 1 - \varepsilon$ for $\varepsilon \in (0, 1]$ and define

$$c_\varepsilon^*(x,\xi) = \widehat{\xi}(x) \frac{f_1(x,\xi)^{1-\varepsilon} - f_1(x,\xi)}{\varepsilon} + \xi(x) \frac{f_0(x,\xi)^{1-\varepsilon} - f_0(x,\xi)}{\varepsilon}.$$

Then the $(1-\varepsilon)$-voter rates may be written as

$$c^{(1-\varepsilon)}(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \varepsilon\, c^*_\varepsilon(x,\xi). \tag{3.36}$$

Recall from (1.20) that for $1 \le \ell \le |\mathcal{N}|$, $r_\ell = (\ell/|\mathcal{N}|)\log(|\mathcal{N}|/\ell)$, and $r_0 = 0$, and let

$$r^\varepsilon_\ell = \frac{(\ell/|\mathcal{N}|)^{1-\varepsilon} - (\ell/|\mathcal{N}|)}{\varepsilon}, \text{ for } \ell \in \{0,\dots,|\mathcal{N}|\}. \tag{3.37}$$

Next, define $g^\varepsilon_i : \{0,1\}^{\mathcal{N}} \to [0,\infty)$ for $i = 0,1$ by

$$g^\varepsilon_i(\xi) = r^\varepsilon_\ell \quad \text{if} \quad \sum_{y\in\mathcal{N}} 1\{\xi(y) = i\} = \ell,$$

and g_i in the same way, but with no superscript ε's. Then we can write

$$\begin{aligned} c^*_\varepsilon(x,\xi) &= \sum_{\ell=1}^{|\mathcal{N}|} r^\varepsilon_l \Big(\widehat{\xi}(x)1\{n_1(x,\xi)=\ell\} + \xi(x)1\{n_0(x,\xi)=\ell\}\Big) \\ &= \widehat{\xi}(x) g^\varepsilon_1(\xi|_{x+\mathcal{N}}) + \xi(x) g^\varepsilon_0(\xi|_{x+\mathcal{N}}). \end{aligned} \tag{3.38}$$

A Taylor series expansion shows that for $0 < u < 1$, and $0 < \varepsilon < 1$,

$$\frac{u^{1-\varepsilon} - u}{\varepsilon} = u\log(1/u) + u\varepsilon \sum_{m=2}^\infty \frac{\varepsilon^{m-2}(\log 1/u)^m}{m!}$$

and thus

$$0 \le \frac{u^{1-\varepsilon} - u}{\varepsilon} - u\log(1/u) \le u\varepsilon \sum_{m=0}^\infty \frac{(\log 1/u)^m}{m!} = \varepsilon. \tag{3.39}$$

The above shows that for $\ell \in \{0,\dots,|\mathcal{N}|\}$,

$$|r^\varepsilon_\ell - r_\ell| \le \varepsilon. \tag{3.40}$$

and (3.40) implies

$$\|g^\varepsilon_i - g_i\|_\infty \le \varepsilon.$$

As $\mathbf{0}$ and $\mathbf{1}$ are traps for q-voter models, we have verified that the collection of q-voter models, $\{c^{(1-\varepsilon)} : \varepsilon \in (0,1]\}$ is a symmetric voter model perturbation with finite range in $\mathcal{N}$, non-negative g^ε_i, and $r_0 = 1$, $c_g = 1$ in (1.14) (even if $d = 2$). As a final note here, using the fact that $\sup_{0<u\le 1} u\log(1/u) = 1/e$, it follows from (3.39) that if $0 < \varepsilon < 1 - 1/e$, then by (3.40),

$$0 < r^\varepsilon_\ell \le 1 \text{ for } 1 \le \ell \le n-1. \tag{3.41}$$

Consider now the rescaled quantities when $d = 2$. We have from the above that

$$r^{N,s}(A) = g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A}) = r^{\varepsilon_N}_{|A|} \to r_{|A|} = r^s(A), \tag{3.42}$$

agreeing with (1.21). The symmetry of $c^{(1-\varepsilon)}$ puts us in the setting of Remark 3.6 and shows $c^{N,a} = c^a = r^{N,a} = r^a = 0$ by (3.25). Therefore,

$$\Theta_3 = \Theta \text{ and } \Theta_2 = 0. \tag{3.43}$$

The first was already noted in (1.22), where the definition of Θ is given.

Example 3.9. (Lotka-Volterra Models). The 2-dimensional Lotka-Volterra model of Neuhauser-Pacala [29] with parameters $(\alpha_0, \alpha_1) \in (0,1)^2$ is the spin-flip system with rate function (see [9])

$$c^{\mathsf{lv}}(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \widehat{\xi}(x)(\alpha_0 - 1)f_1(x,\xi)^2 + \xi(x)(\alpha_1 - 1)f_0(x,\xi)^2.$$

Here both the voter model kernel and the notation f_i use the kernel $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$ for some neighbourhood $\mathcal{N}$. If $\alpha_1 \vee \alpha_2 \geq \frac{1}{2}$, c^{lv} is monotone (see Section 1 of [11]). In the diagonal case $\alpha_1 = \alpha_2$, c^{lv} is good cancellative (see, for example, the Proof of Theorem 1.1 in Section 6 of [13]). As in Section 1 of [9], for $\varepsilon \in (0,1)$ one sets

$$\alpha_i = 1 - \varepsilon + \beta_i^{(\varepsilon)}(\log(1/\varepsilon))^{-2}\varepsilon, \text{ where } \lim_{\varepsilon\downarrow 0}\beta_i^{(\varepsilon)} = \beta_i \in \mathbb{R}, \quad i = 0, 1,$$

so that our voter model perturbation rates are given by

$$\begin{aligned} c_\varepsilon(x,\xi) &= c^{\mathsf{vm}}(x,\xi) + \varepsilon g_1^\varepsilon(\xi|_{x+\mathcal{N}})\widehat{\xi}(x) + \varepsilon g_0^\varepsilon(\xi|_{x+\mathcal{N}})\xi(x), \\ g_i^\varepsilon(\xi|_{\mathcal{N}}) &= \big(-1 + (\log 1/\varepsilon)^{-2}\beta_{1-i}^{(\varepsilon)}\big)f_i(0,\xi)^2, \quad i = 0, 1. \end{aligned}$$

If $g_i(\xi|_{\mathcal{N}}) = -f_i(0,\xi)^2$, clearly

$$\|g_i^\varepsilon - g_i\|_\infty \leq C(\log(1/\varepsilon))^{-2} \to 0 \quad \text{for } i = 0, 1.$$

Obviously both $\mathbf{0}$ and $\mathbf{1}$ are traps for each ε, and we also have

$$\begin{aligned} \lim_{\varepsilon\downarrow 0}(\log(1/\varepsilon))^2(g_1^\varepsilon(\xi) - g_0^\varepsilon(\widehat{\xi})) &= \lim_{\varepsilon\downarrow 0}(\beta_0^{(\varepsilon)} - \beta_1^{(\varepsilon)})f_1(0,\xi)^2 \\ &= (\beta_0 - \beta_1)f_1(0,\xi)^2 \\ &:= g^a(\xi). \end{aligned}$$

Therefore $\{c_\varepsilon : \varepsilon \in (0,1)\}$ is an asymptotically symmetric finite range voter model perturbation. We have

$$\beta_i^N := \Big(\frac{\log N}{\log(1/\varepsilon_N)}\Big)^2\beta_i^{(\varepsilon_N)} \to \beta_i \text{ as } N \to \infty,$$

and so, for $\emptyset \neq A \subset \mathcal{N}$,

$$\begin{aligned} r^{N,s}(A) = g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A}) &= (-1 + (\log N)^{-2}\beta_1^N)\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2 \\ &\to -\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2 = r^s(A), \end{aligned} \tag{3.44}$$

and

$$\begin{aligned} r^{N,a}(A) &= (\log N)^2[g_1^{\varepsilon_N}(1_A) - g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A})] \\ &= (\beta_0^N - \beta_1^N)f_1(0,1_A)^2 \\ &\to (\beta_0 - \beta_1)\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2 = r^a(A). \end{aligned}$$

Therefore in this case we have

$$\Theta_2 = \Theta_2^{\mathsf{lv}} := (\beta_0 - \beta_1)\sum_{\emptyset\neq A\subset\mathcal{N}}\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2 K_2(A, \bar{\mathcal{N}} \setminus A), \tag{3.45}$$

and

$$\Theta_3 = \Theta_3^{\mathsf{lv}} := \sum_{\emptyset\neq A\subset\mathcal{N}}\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2(\Theta^-(A) - \Theta^+(A)). \tag{3.46}$$

Example 3.10. (Affine Voter Models) The 2-dimensional threshold voter model rate function, introduced in [5], and corresponding to $q = 0$ in Example 3.8, is

$$c^{\mathsf{tv}}(x,\xi) = 1\{\xi(x+y) \neq \xi(x) \text{ for some } y \in \mathcal{N}\}$$

for a neighbourhood $\mathcal{N}$. The affine voter model with parameter $\alpha \in [0,1]$, introduced in [31], is the spin-flip system with rate function

$$c^{\mathsf{av}}(x,\xi) = \alpha c^{\mathsf{vm}}(x,\xi) + (1-\alpha)c^{\mathsf{tv}}(x,\xi).$$

Here the voter model kernel is $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, and $\mathbf{0}$ and $\mathbf{1}$ are traps. The fact that c^{tv} is cancellative was noted in Section 2 of [5]. It is easy to check that a convex combination of cancellative rate functions is cancellative, and hence c^{av} is cancellative, in fact good cancellative for $\alpha < 1$. Monotonicity of c^{tv}, and hence of c^{av} is clear. Letting $\varepsilon = 1-\alpha$,

$$c^{\mathsf{av}}(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \varepsilon(c^{\mathsf{tv}}(x,\xi) - c^{\mathsf{vm}}(x,\xi))$$

so that our voter model perturbation rates are given by

$$\begin{aligned} c_\varepsilon(x,\xi) &= c^{\mathsf{vm}}(x,\xi) + \varepsilon\big(\widehat{\xi}(x)g_1^\varepsilon(\xi|_{x+\mathcal{N}}) + \xi(x)g_0^\varepsilon(\xi|_{x+\mathcal{N}})\big), \\ g_i^\varepsilon(\xi|_{\mathcal{N}}) &= -f_i(0,\xi) + 1\{n_i(0,\xi) \geq 1\} \quad i = 0,1. \end{aligned}$$

Since $g_i = g_i^\varepsilon$ does not depend on ε, (1.15) holds. Obviously each c_ε is symmetric, so $r^{N,a} = r^a = g^a \equiv 0$. Therefore $\{c_\varepsilon : \varepsilon \in (0,1)\}$ is an asymptotically symmetric finite range voter model perturbation such that for $A \neq \emptyset$,

$$r^{N,s}(A) = r^s(A) = g_0(1_{\mathcal{N}\setminus A}) = -\frac{|A|}{|\mathcal{N}|} + 1,$$

$\Theta_2 = 0$, and if we use (3.34) and then (3.35), we get

$$\Theta_3 = \sum_{\emptyset\neq A\subset\mathcal{N}} \Big(1 - \frac{|A|}{|\mathcal{N}|}\Big)(\Theta^+(A) - \Theta^-(A)) = \sum_{\emptyset\neq A\subset\mathcal{N}} (\Theta^+(A) - \Theta^-(A)) = \kappa \,:= \Theta_3^{\mathsf{av}}. \quad (3.47)$$

Example 3.11. (Geometric Voter Models) The 2-dimensional geometric voter model with rate function, introduced in [5], is

$$c^{\mathsf{gv}}(x,\xi) = \frac{1-\theta^j}{1-\theta^{|\mathcal{N}|}} \text{ if } \sum_{y\in\mathcal{N}} 1\{\xi(x+y) \neq \xi(x)\} = j. \quad (3.48)$$

Here $\mathcal{N}$ is a neighbourhood, $0 \leq \theta < 1$, and $\mathbf{0}$ and $\mathbf{1}$ are clearly traps. As θ ranges from 0 to 1, these dynamics range from the threshold voter model to the voter model. The fact that c^{gv} is cancellative was proved in Section 2 of [5], and it follows that c^{gv} is then good cancellative. Monotonicity of c^{gv} is again elementary. By (7.3) in [13], taking $\varepsilon = 1-\theta$ (instead of ε^2),

$$c^{\mathsf{gv}}(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \varepsilon\frac{|\mathcal{N}|}{2}f_0(x,\xi)f_1(x,\xi) + O(\varepsilon^2) \text{ as } \varepsilon \to 0. \quad (3.49)$$

Here, c^{vm} and the densities f_i use the kernel $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, and the $O(\varepsilon^2)$ term is uniform in x, ξ. Thus, there are g_i^ε such that

$$c_\varepsilon(x,\xi) = c^{\mathsf{vm}}(x,\xi) + \varepsilon g_1^\varepsilon(\xi|_{x+\mathcal{N}})\widehat{\xi}(x) + \varepsilon g_0^\varepsilon(\xi|_{x+\mathcal{N}})\xi(x),$$

where if $g(x,\xi) = (|\mathcal{N}|/2)f_0(x,\xi)f_1(x,\xi)$, then $\|g_i^\varepsilon - g\| \leq C|\varepsilon|$. Clearly c^{gv} is symmetric and so by Remark 3.6 $g^a = r^a = 0$. Therefore $\{c_\varepsilon, 0 \leq \varepsilon < 1\}$ is an asymptotically symmetric voter model perturbation such that for $A \neq \emptyset$,

$$r^s(A) = g(1_{\mathcal{N}\setminus A}) = \frac{|A|(|\mathcal{N}| - |A|)}{2|\mathcal{N}|}.$$

Thus $\Theta_2 = 0$ and, again using (3.34), we have

$$\Theta_3 = \frac{1}{2|\mathcal{N}|}\sum_{\emptyset\neq A\subset\mathcal{N}}|A|(|\mathcal{N}|-|A|)(\Theta^+(A)-\Theta^-(A)) = \frac{|\mathcal{N}|}{2}\sum_{\emptyset\neq A\subset\mathcal{N}}\Big(\frac{|A|}{|\mathcal{N}|}\Big)^2(\Theta^-(A)-\Theta^+(A)) \tag{3.50}$$

$$:= \Theta_3^{\mathtt{gv}} = \frac{|\mathcal{N}|}{2}\Theta_3^{\mathtt{lv}}.$$

3.3 Proof of Theorem 1.1 for $d \geq 3$

We assume $\mathcal{N}$ is as in Theorem 1.1 for $d \geq 3$. Let $\langle\quad\rangle_u$ denote expectation with respect to the voter model equilibrium with density $u \in [0,1]$ and define symmetric rates

$$f(u) = \langle (1-\xi(0))c^*(0,\xi) - \xi(0)c^*(0,\xi)\rangle_u \text{ for } u \in [0,1], \tag{3.51}$$

where c^* is as in (3.38) but with r_ℓ in place of r_ℓ^ε. As noted in Section 1 of [7] f will be a polynomial in u of degree at most $|\mathcal{N}|+1$. As we have $d \geq 3$, Theorem 1.2 of [13] (as strengthened in Remark 1.10) shows that the conclusion of Theorem 1.1 will hold for $q < 1$ and sufficiently close to 1, providing the following hold for some $\varepsilon_0 \in (0,1)$:

(1) For $0 < \varepsilon < \varepsilon_0$, $c^{(1-\varepsilon)}(x,\xi)$ is a rate function of a cancellative process (as in (3.1)).

(2) $\{c^{(1-\varepsilon)}(x,\xi) : 0 < \varepsilon < \varepsilon_0\}$ is a finite range voter model perturbation.

(3) $f'(0) > 0$.

Here we have ignored one condition from Theorem 1.2 of [13] (condition (1.2) there) as the required exponential tail bound is trivially true for our finite neighbourhood setting. The first two conditions have been established in Sections 3.1 (Lemma 3.3) and 3.2, respectively. Finally (3) follows from Theorem 1.2 of [4], where it is shown for any neighbourhood $\mathcal{N}$. □

Remark 3.12. The proof of $f'(0) > 0$ from [4] is stated for $d = 3$ but holds equally well for $d \geq 3$. In fact it gives a stronger representation for f which implies $f'(0) > 0$. Alternatively see Proposition 5.9 for a more direct proof.

4 A general complete convergence theorem in two dimensions

Fix an initial state $\xi_0 \in \{0,1\}^{\mathbb{Z}^2}$. For $\xi,\eta \in \{0,1\}^{\mathbb{Z}^2}$. We first extend the stochastic differential equation (SDE) construction of spin-flip systems and coupling with killed processes from Section 2 of [11] to the setting where $|\xi_0|$ may be infinite. Let $\{N^{x,i} : x \in \mathbb{Z}^2, i = 0,1\}$ be independent Poisson point processes on $\mathbb{R}_+^2$ with rate $ds \times du$. For $R \subset \mathbb{R}^2$ and $T \geq 0$,

$$\mathcal{G}([0,T]\times R) \text{ is the } \sigma\text{-field generated by } \{N^{x,i}|_{[0,T]\times\mathbb{R}} : x \in R, i = 0,1\}. \tag{4.1}$$

Consider a rate function $c : \mathbb{Z}^2 \times \{0,1\}^{\mathbb{Z}^2} \to [0,\infty)$, which is bounded continuous. Let $\hat{c}(x,\xi) = c(x,\widehat{\xi})$ be the rate function for the evolution of the 0's. We assume

$$\sum_x \big(c(x,\xi) + \hat{c}(x,\xi)\big) \leq C|\xi|, \tag{4.2}$$

and

$$\sup_{x\in\mathbb{Z}^2}\sum_{u\in\mathbb{Z}^2}\sup_\xi |c(x,\xi) - c(x,\xi^{(u)})| \leq C, \tag{4.3}$$

where

$$\xi^{(u)}(x) = 1(x \neq u)\xi(x) + 1(x = u)\widehat{\xi}(x).$$

Under (4.3) and boundedness and continuity of c, there is a unique Feller process ξ_t taking values in $\{0,1\}^{\mathbb{Z}^2}$ associated with the rate function c (see Theorem B3 in [25]). Note that the above conditions hold for c iff they hold for $\hat{c}$.

Remark 4.1. It is easy to check all of the above conditions are satisfied by c_ε, if $\{c_\varepsilon : 0 < \varepsilon \le \varepsilon_0\}$ are the rates of a finite range voter model perturbation. See Corollary 2.4 of [11] for (4.2) (without the $\hat{c}$ term) and (4.3). Condition (4.2) follows easily for $\hat{c}$ from the fact that $\mathbf{1}$ is a trap for our finite range voter model perturbations in $d = 2$ (just as it followed for c using the fact that $\mathbf{0}$ is a trap). Note that under our boundedness assumption on c, (4.2) for c alone is equivalent to condition (2.1) in [11].

For $\xi_0 \in \{0,1\}^{\mathbb{Z}^2}$ consider the SDE

$$\begin{aligned}\xi_t(x) = \xi_0(x) + \int_0^t \int (1-\xi_{s-}(x))1(u \le c(x,\xi_{s-}))N^{x,0}(ds,du)\\ - \int_0^t \int \xi_{s-}(x)1(u \le c(x,\xi_{s-}))N^{x,1}(ds,du), \quad t \ge 0, x \in \mathbb{Z}^2.\end{aligned} \tag{4.4}$$

If $|\xi_0| < \infty$, (4.4) has a pathwise unique solution $\xi_t = \xi_t[\xi_0]$ which has the same law as the above Feller process with initial condition ξ_0. For this see Proposition 2.1(a) of [11] and note, that monotonicity of c is not needed in Proposition 2.1(c). Now fix $M_0 \in \mathbb{N}$, set $I' = (-M_0, M_0)^2$, and let $\underline{c}(x,\xi) = 1(x \in I')c(x,\xi)$. Clearly $\underline{c}$ also satisfies all the hypotheses we have imposed on c. For our given initial condition ξ_0, let $\underline{\xi}_0(x) = 1(x \in I')\xi_0(x)$. Then $|\underline{\xi}_0| < \infty$ (even if $|\xi_0| = \infty$) and so there is a unique solution, $\underline{\xi}_t = \underline{\xi}_t[\underline{\xi}_0, I']$, of

$$\begin{aligned}\underline{\xi}_t(x) = \underline{\xi}_0(x) + \int_0^t \int (1-\underline{\xi}_{s-}(x))1(u \le \underline{c}(x,\underline{\xi}_{s-}))N^{x,0}(ds,du)\\ - \int_0^t \int \underline{\xi}_{s-}(x)1(u \le \underline{c}(x,\underline{\xi}_{s-}))N^{x,1}(ds,du), \quad t \ge 0, x \in \mathbb{Z}^2.\end{aligned} \tag{4.5}$$

Note that $\underline{\xi}_t(x) = 0$ for all $t \ge 0$ and all $x \notin I'$. If, in addition, c is monotone, then by Proposition 2.1(b) of [11],

$$\text{if } \xi_\cdot \text{ satisfies (4.4) where } |\xi_0| < \infty, \text{ then } \underline{\xi}_t \le \xi_t \ \forall t \ge 0. \tag{4.6}$$

We now show this continues to hold even if $|\xi_0| = \infty$. **Unless otherwise indicated, monotonicity of the rate function, c, is assumed in the rest of this section.** Note that $\hat{c}$ is monotone iff c is. Therefore our hypotheses hold for c iff they hold for $\hat{c}$.

Lemma 4.2. *If $\xi_\cdot$ satisfies* (4.4) *then $\underline{\xi}_t \le \xi_t$ for all $t \ge 0$ a.s.*

Proof. By (4.6), we may focus on $|\xi_0| = \infty$. Define

$$\Lambda_t = \sum_{x \in I'} \int_0^t \int_0^{\|c\|_\infty} (N^{x,0} + N^{x,1})(ds,du).$$

Then Λ is a Poisson process with rate $2|I'|\|c\|_\infty < \infty$ and so has a sequence of jump times $0 < T_1 < T_2 < \dots$ increasing to infinity. It suffices to prove that $\underline{\xi}_t \le \xi_t$ for $t \in [0, T_n]$ for all n, which we prove by induction. Assume the inequality up to, and including, time T_n. On (T_n, T_{n+1}), $\underline{\xi}_t(x) = \underline{\xi}_{T_n}(x) \le \xi_{T_n}(x) = \xi_t(x)$ for all $x \in I'$ since the jump times of these coordinates are clearly included in the jump times of Λ. For $x \notin I'$, $\underline{\xi}_t(x) = 0 \le \xi_t(x)$ for all t, including those in (T_n, T_{n+1}). At T_{n+1} one considers the unique $x \in I'$ for which $N^{x,i}$ jumps at T_{n+1} for some i and uses the monotonicity of c to show that in each of the two cases $\Delta\xi_{T_{n+1}}(x) < 0$ or $\Delta\underline{\xi}_{T_{n+1}}(x) > 0$, one has $\underline{\xi}_{T_{n+1}}(x) \le \xi_{T_{n+1}}(x)$. This is done just as in the proof of Proposition 2.1(b) in [11]. The remaining cases trivially lead to the same conclusion. This establishes the induction step, and the $n = 1$ step is handled in exactly the same way. □

If $\xi_\cdot$ satisfies (4.4), then $\widehat{\xi}_\cdot$ satisfies

$$\widehat{\xi}_t(x) = \widehat{\xi}_0(x) + \int_0^t \int (1 - \widehat{\xi}_{s-}(x))1(u \le \hat{c}(x, \widehat{\xi}_{s-}))N^{x,1}(ds, du) \\ - \int_0^t \int \widehat{\xi}_{s-}(x)1(u \le \hat{c}(x, \widehat{\xi}_{s-}))N^{x,0}(ds, du), \quad t \ge 0, x \in \mathbb{Z}^2. \quad (4.7)$$

Note that c has been replaced with $\hat{c}$ and the roles of $N^{x,0}$ and $N^{x,1}$ have been reversed from that in (4.4). We set $\hat{\underline{\xi}}_0(x) = \widehat{\xi}_0(x)1(x \in I')$ and define $\hat{\underline{\xi}}_t = \hat{\underline{\xi}}_t[\hat{\underline{\xi}}_0, I']$ as the unique solution (by Proposition 2.1(a) of [11] because $\hat{c}$ satisfies the same hypotheses as c) to

$$\hat{\underline{\xi}}_t(x) = \hat{\underline{\xi}}_0(x) + \int_0^t \int (1 - \hat{\underline{\xi}}_{s-}(x))1(u \le \hat{\underline{c}}(x, \hat{\underline{\xi}}_{s-}))N^{x,1}(ds, du) \\ - \int_0^t \int \hat{\underline{\xi}}_{s-}(x)1(u \le \hat{\underline{c}}(x, \hat{\underline{\xi}}_{s-}))N^{x,0}(ds, du), \quad t \ge 0, x \in \mathbb{Z}^2. \quad (4.8)$$

So in the notation $\hat{\underline{\xi}}$ we effectively take the hat first and then do the killing.

If $M \in \mathbb{N}$ and $\xi_0 \in \{0,1\}^{\mathbb{Z}^2}$, we may set $\underline{\xi}_0(x) = \xi_0^M(x) := \xi_0(x)1(x \in (-M, M)^2)$, and denote the unique solution to (4.4) with this initial state by ξ_t^M.

Proposition 4.3. *(a) As $M \to \infty$, $\xi_t^M(x) \uparrow \xi_t^\infty(x)$ for all $x \in \mathbb{Z}^2$ and $t \ge 0$ a.s. Moreover ξ_t^∞ is the unique in law Feller process with rates $c(x, \xi)$ starting at ξ_0. If $|\xi_0| < \infty$ then $\xi_t^\infty = \xi_t[\xi_0]$, the unique solution of(4.4).*
(b) We have $\underline{\xi}_t(x) \le \xi_t^\infty(x)$, and $\hat{\underline{\xi}}_t(x) \le \widehat{\xi_t^\infty}(x)$ for all $x \in \mathbb{Z}^2$ $t \ge 0$ a.s.
(c) If c is symmetric, then $P(\hat{\underline{\xi}} \in \cdot) = P_{\hat{\underline{\xi}}_0}(\underline{\xi} \in \cdot)$, where the right-hand side is the law of $\underline{\xi}_\cdot$ with initial state $\hat{\underline{\xi}}_0$, that is, the law of $\underline{\xi}_\cdot[\hat{\underline{\xi}}_0, I']$.

Proof. (a) By Proposition 2.1(b) of [11], $\xi_t^M(x)$ increases in M for all (t, x) a.s. and so we can define ξ^∞ as this a.s. limit. By (c) of the same Proposition the law of ξ^M is that of the unique Feller process with rates c and initial condition ξ_0^M. Theorem B3 of [25] allows us to apply Theorem 5.2 of [24] to conclude that the martingale problem associated with the rates c is well-posed, and then Proposition 6.5 of [24] gives continuity of the laws in the initial condition. This implies that ξ^∞ has the required law. If $|\xi_0| < \infty$ and $\xi_t = \xi_t[\xi_0]$ is the unique solution of (4.4), then monotonicity in the initial condition from Proposition 2.1(b) of [11] shows that $\xi_t^M \le \xi_t$ for all $t \ge 0$ a.s., and so taking limits we get $\xi_t^\infty \le \xi_t$ for all $t \ge 0$ a.s. Since $\xi_\cdot$ and $\xi_\cdot^\infty$ have the same law, they must be identical.
(b) Let $M \ge M_0$, so that $\underline{\xi}_0^M := \xi_0^M(x)1(x \in I') = \underline{\xi}_0(x)$ and therefore $\underline{\xi}_t = \underline{\xi}_t^M := \underline{\xi}_t(\underline{\xi}_0^M, I')$. By (4.6), w.p.1 for any $t \ge 0$,

$$\underline{\xi}_t = \underline{\xi}_t^M \le \xi_t^M \le \xi_t^\infty.$$

If $\hat{\underline{\xi}}_t^M = \hat{\underline{\xi}}_t(\hat{\xi}_0^M, I')$, then, just as above with $\hat{c}$ in place of c, $\hat{\underline{\xi}}_t = \hat{\underline{\xi}}_t^M$. The process $\widehat{\xi}_t^M = 1 - \xi_t^M$ satisfies (4.7), and so we may apply Lemma 4.2 with $\hat{c}$ in place of c and the roles of $N^{x,0}$ and $N^{x,1}$ reversed and so conclude that

$$\hat{\underline{\xi}}_t = \hat{\underline{\xi}}_t^M \le 1 - \xi_t^M \to \widehat{\xi_t^\infty} \text{ as } M \to \infty \quad \forall t \ge 0 \text{ a.s.},$$

and so deduce the second inequality in (b).

(c) Under symmetry of c, $\underline{\hat{\xi}}$ is the unique solution of (4.8) with $\hat{c} = c$, and so has the same law as $\underline{\xi}[\hat{\underline{\xi}}_0, I']$, the unique solution of (4.5) with initial condition $\hat{\underline{\xi}}_0$, because $(N^{x,0}, N^{x,1})$ is equal in law to $(N^{x,1}, N^{x,0})$. □

In view of (a) of the above we will denote ξ_t^∞ by $\xi_t[\xi_0]$ as it agrees with our earlier definition for $|\xi_0| < \infty$. The reader will note however, we have side-stepped the general question of pathwise existence and uniqueness of solutions to (4.4) when $|\xi_0| = \infty$. We believe this to be the case by uniqueness of the martingale problem and monotonicity, but will not need it.

(a) and (b) show that for any initial state ξ_0, and corresponding $\underline{\xi}_0$ and $\hat{\underline{\xi}}_0$, we may construct $(\xi, \underline{\xi}, \hat{\underline{\xi}})$ on the same space such that

$$\underline{\xi}_t \le \xi_t \text{ and } \hat{\underline{\xi}}_t \le \widehat{\xi}(t) \ \ \forall t \ge 0 \ a.s. \tag{4.9}$$

Here $\underline{\xi}$ evolves according to rate c with killing outside I', and $\underline{\hat{\xi}}$ evolves according to rate $\hat{c}$ with the same killing, while ξ and $\widehat{\xi}$ evolve according the rates c and $\hat{c}$, respectively with no killing. Under symmetry, the dynamics of the two killed processes are the same but of course the initial conditions differ. In the symmetric case this point seems to be made implicitly in the proof of Theorem 1.1 in Section 6 of [13], but perhaps warrants the explicit construction given above.

Turning to the complete convergence theorem, we first present an abstract complete convergence in two dimensions essentially taken from [13]. If $A \subset \mathbb{Z}^2$, $x_0 \in \mathbb{Z}^2$ and $\xi \in \{0,1\}^{\mathbb{Z}^2}$, let

$$A(x_0, \xi) = \{y \in A : \xi(y) = 1, \xi(y + x_0) = 0\}.$$

Monotonicity is not required for our first abstract complete convergence theorem.

Theorem 4.4. *Assume for $0 < \varepsilon \le \varepsilon_0$, $\xi^{[\varepsilon]}$ is a cancellative finite range voter model perturbation with rate function $c_\varepsilon(x, \xi)$. Assume also that for each ε,*

$$\exists x_0 \in \mathbb{Z}^2 \text{ so that if } |\widehat{\xi_0^{[\varepsilon]}}| = \infty \text{ then } \lim_{K\to\infty} \sup_{A\subset\mathbb{Z}^2, |A|\ge K} \lim_{t\to\infty} P_{\xi_0^{[\varepsilon]}}(|\xi_t^{[\varepsilon]}| > 0,\ A(x_0, \xi_t^{[\varepsilon]}) = \emptyset) = 0, \tag{4.10}$$

and

$$\limsup_{t\to\infty} P_{\delta_0}(\xi_t^{[\varepsilon]}(0) = 1) > 0. \tag{4.11}$$

There is an $\varepsilon_1 > 0$ such that for $\varepsilon \in (0, \varepsilon_1)$, there is a translation invariant symmetric stationary distribution $\nu_{1/2}^\varepsilon$ with density $1/2$, satisfying the coexistence property, such that for all initial $\xi_0^{[\varepsilon]}$,

$$\xi_t^{[\varepsilon]} \Rightarrow \beta_0(\xi_0^{[\varepsilon]})\delta_{\mathbf{0}} + \beta_\infty(\xi_0^{[\varepsilon]})\nu_{1/2}^\varepsilon + \beta_1(\xi_0^{[\varepsilon]})\delta_{\mathbf{1}} \quad \text{as } t \to \infty. \tag{4.12}$$

Proof. We first show that for $\varepsilon < \varepsilon_0$, $\xi^{[\varepsilon]}$ is a good cancellative process (as defined in Section 3.1). Assume not. By Remark 3.1, up to a constant time change, $\xi^{[\varepsilon]}$ is a voter model with kernel $q_0(y) = \beta_0(\{y\})$. Therefore starting at $\xi_0^{[\varepsilon]} = \delta_0$, $\xi_t^{[\varepsilon]} = \mathbf{0}$ for t large a.s. (e.g. see Proposition V.4.1(b) of [24]), and this contradicts (4.11), completing the proof.

Under the above hypotheses, Remark 4, Corollary 3.3, and Lemma 4.2 of [13] show that for some $\varepsilon_1 > 0$, the hypotheses of Proposition 4.1 of [13] hold for $\varepsilon \in (0, \varepsilon_1)$. (Note that the *good* cancellative property is needed to apply Corollary 3.3.) That Proposition implies the stationary measure $\nu_{1/2}^\varepsilon$ (from Section 3.1) satisfies the coexistence property and

$$\text{if } |\widehat{\xi_0^{[\varepsilon]}}| = \infty, \text{ then } \xi_t^{[\varepsilon]} \Rightarrow \beta_0(\xi_0^{[\varepsilon]})\delta_{\mathbf{0}} + \beta_\infty(\xi_0^{[\varepsilon]})\nu_{1/2}^\varepsilon + \beta_1(\xi_0^{[\varepsilon]})\delta_{\mathbf{1}} \text{ as } t \to \infty, \tag{4.13}$$

where we have also used (1.3). The fact that $\mathbf{0}$ is a trap for $\xi^{[\varepsilon]}$ (by (1.16)) implies $\xi^{[\varepsilon]}$ is symmetric by the equivalence of (3.3) and (3.4) noted in Section 3.1. Therefore $\nu^{\varepsilon}_{1/2}$ is also a symmetric law ($\nu^{\varepsilon}_{1/2}(\widehat{\xi} \in \cdot) = \nu^{\varepsilon}_{1/2}(\cdot)$) by the above convergence. By symmetry the conclusion of (4.13) also holds if $|\xi^{[\varepsilon]}_0| = \infty$ and so the proof of (4.12) is complete. □

Remark 4.5. As noted in the above proof, a cancellative finite range voter model perturbation is symmetric.

The key condition in the above is (4.10) which will imply that a pair of nearby sites with opposite type can be found in sufficiently large sets for large t. Such pairs are to be expected if there is to be complete convergence with coexistence.

Assume now that for some $\varepsilon_0 > 0$:

$$\text{For } 0 < \varepsilon \leq \varepsilon_0,\ \xi^{[\varepsilon]} \text{ is a cancellative and monotone finite range voter model perturbation in } \mathbb{Z}^2. \tag{4.14}$$

By Remarks 4.1 and 4.5 we may apply Proposition 4.3, and so for any initial condition $\xi_0 \in \{0,1\}^{\mathbb{Z}^2}$ construct $\xi^{[\varepsilon]}$, $\underline{\xi}^{[\varepsilon]}$, $\hat{\underline{\xi}}^{[\varepsilon]}$ as solutions of (4.4), (4.5) and (4.8), respectively, with $\xi^{[\varepsilon]}_0 = \xi_0$, all on a common probability space such that

$$\underline{\xi}^{[\varepsilon]}_t \leq \xi^{[\varepsilon]}_t \text{ and } \hat{\underline{\xi}}^{[\varepsilon]}_t \leq \widehat{\xi^{[\varepsilon]}_t} \quad \forall t \geq 0 \ a.s. \tag{4.15}$$

We now assume that $M_0 = KL$ for natural numbers K, L chosen below. As in Section 3.2 we often use $N \geq N(\varepsilon_0) > e^3$ satisfying $\varepsilon = \varepsilon_N := \frac{(\log N)^3}{N}$ as our fundamental parameter. Let

$$\underline{\xi}^N_t(x) = \underline{\xi}^{[\varepsilon_N]}_{Nt}(\sqrt{N}x),\ x \in S_N, \tag{4.16}$$

and

$$\underline{X}^N_t = \frac{\log N}{N} \sum_{x \in S_N} \underline{\xi}^N_t(x)\delta_x, \text{ and } \hat{\underline{X}}^N_t = \frac{\log N}{N} \sum_{x \in S_N} \hat{\underline{\xi}}^N_t(x)\delta_x. \tag{4.17}$$

The next condition is the key to ensure the survival of our oriented percolation process. It specifies the values of K, L which are used above to define our killed particle systems through $M_0 = KL$. Note that $\underline{X}^N_t(\mathbf{1}) = \underline{X}^N_t((-M_0, M_0)^2) < \infty$.

$$\begin{aligned}&\text{There are } T' > 1,\ K, J' \in \mathbb{N} \text{ with } K > 2, \text{ and } L' > 3, \text{ so that if}\\ &0 < \varepsilon \leq \varepsilon_0, \text{ and } I_{\pm e_i} = \pm 2L'e_i + [-L'+1, L'-1]^2, \text{ then for } L = \lfloor \sqrt{N}L' \rfloor,\\ &\underline{X}^N_0([-L', L']^2) \geq J' \text{ implies } P(\underline{X}^N_{T'}(I_e) \geq J' \text{ for all } e \in \{\pm e_i, i = 1, 2\}) \geq 1 - 6^{-5(2K+1)^3}.\end{aligned} \tag{4.18}$$

Recall that if the conclusion of Theorem 4.4 holds we say for $0 < \varepsilon < \varepsilon_1$ the complete convergence theorem with coexistence (CCT) holds for $\xi^{[\varepsilon]}$. We use this terminology going forward.

Theorem 4.6. *Assume for $0 < \varepsilon \leq \varepsilon_0$, $\xi^{[\varepsilon]}$ is a cancellative and monotone finite range voter model perturbation in $\mathbb{Z}^2$ satisfying (4.18). There is an $\varepsilon_1 > 0$ such that for $\varepsilon \in (0, \varepsilon_1)$, the complete convergence theorem with coexistence (CCT) holds for $\xi^{[\varepsilon]}$.*

Proof. We follow the general approach used in Section 6 of [13] for the $d = 2$ Lotka-Volterra model. From Theorem 4.4 it suffices to establish (4.10) and (4.11), as well as the fact that $\beta_\infty(\xi^{[\varepsilon]}_0) > 0$ for initial conditions distinct from $\mathbf{0}$ and $\mathbf{1}$ (we have suppressed the dependence on ε in β_∞). Remark 4.5 shows that $\xi^{[\varepsilon]}$ is symmetric. Therefore, we may use Proposition 4.3(c) to see that (4.18) implies that,

$$\begin{aligned}&\underline{X}^N_0([-L', L']^2) \geq J' \text{ and } \hat{\underline{X}}^N_0([-L', L']^2) \geq J' \text{ imply}\\ &\quad P(\underline{X}^N_{T'}(I_{\pm e_i}) \geq J' \text{ and } \hat{\underline{X}}^N_{T'}(I_{\pm e_i}) \geq J' \text{ for } i = 1, 2) \geq 1 - 2 \cdot 6^{-5(2K+1)^3}.\end{aligned} \tag{4.19}$$

To undo the scaling, recall $L = \lfloor \sqrt{N} L' \rfloor$ and let $J = \frac{N}{\log N} J'$, $T = NT'$ and set $\tilde{I}_{\pm e_i} = \pm 2Le_i + [-L, L]^2$, $i = 1, 2$. In order to use Theorem 4.3 of [16] we introduce a set $H \subset \{0,1\}^{\mathbb{Z}^2}$ of "happy" configurations and a good event $G_{\xi_0} = G_{\xi_0}(\varepsilon)$ in our probability space for each initial condition ξ_0. We let

$$H = \{\xi \in \{0,1\}^{\mathbb{Z}^2} : \xi([-L, L]^2) \geq J \text{ and } \widehat{\xi}([-L, L]^2) \geq J\}, \tag{4.20}$$

and

$$G_{\xi_0} = \{\underline{\xi}_T^{[\varepsilon]}(\tilde{I}_{\pm e_i}) \geq J,\ \hat{\underline{\xi}}_T^{[\varepsilon]}(\tilde{I}_{\pm e_i}) \geq J \text{ for } i = 1, 2\}, \tag{4.21}$$

where ξ_0 is the initial condition for $\xi^{[\varepsilon]}$, so that $\underline{\xi}_0^{[\varepsilon]}(x) = 1_{I'}(x)\xi_0(x)$ and $\hat{\underline{\xi}}_0^{[\varepsilon]}(x) = 1_{I'}(x)\widehat{\xi}_0(x)$. Recall here that $I' = (-KL, KL)^2$. For $z \in \mathbb{Z}^2$, $\sigma_z : \{0,1\}^{\mathbb{Z}^2} \to \{0,1\}^{\mathbb{Z}^2}$ is the translation map, $\sigma_z(\xi)(x) = \sigma(x + z)$. Note that:

(i) G_{ξ_0} is $\mathcal{G}(I' \times [0, T])$-measurable for each ξ_0.

(ii) If $\xi_0 \in H$, then on G_{ξ_0}, $\xi_T^{[\varepsilon]} \in \sigma_{2Le}(H)$ for all $e \in \{\pm e_1, \pm e_2\}$.

(iii) For any $\xi_0 \in H$, $P(G_{\xi_0}) \geq 1 - 2 \cdot 6^{-5(2K+1)^3} := 1 - \gamma'$.

Properties (i) and (ii) are clear from the definitions and the orderings in (4.15). Property (iii) follows from (4.19), along with a bit of arithmetic on the rescaled intervals to show $\sqrt{N} I_{\pm e_i} \subset \tilde{I}_{\pm e_i}$, where $N \geq e^3$ is used. Finally another bit of arithmetic shows that $(1 - \gamma')^{1/(2K+1)^3} > 1 - 6^{-4}$, ensuring that (5.17) of [13] is valid. In this way we have established the set-up of Lemma 5.2 of [13] and we can invoke the comparison with $2K$-dependent percolation from Theorem 4.3 of [16], as carried out in Section 5 of [13]. In particular, we may use the proof of Lemma 5.3 of [13] to conclude that (4.10) and (4.11) hold for $\xi^{[\varepsilon]}$. Although the hypotheses of that result require $d \geq 3$ and $f'(0) > 0$ for a solution to a reaction diffusion equation (which is not defined in $d = 2$), those hypotheses are only used to establish the set-up in Lemma 5.2 of [13], which we have just verified directly, essentially using (4.18). The rest of the proof of Lemma 5.3 of [13] only requires arguments for general voter model perturbations and, in particular, uses its branching coalescing dual from Section 2 of [7] to bound some probabilities involving $\xi^{[\varepsilon]}$. n this way the proof of Lemma 5.3 of [13] gives us (4.10), and the proof also gives (this is (5.31) of [13])

$$\inf_{\xi \neq 0} P_\xi(\xi_t^{[\varepsilon]} \neq \mathbf{0}\ \forall t \geq 0) \geq \rho, \tag{4.22}$$

for some explicit $\rho > 0$. After the proof of Lemma 5.3 in [13], (4.11) is derived from (4.22) using the above oriented percolation setting and elementary properties of voter model perturbations, which apply equally well in our setting.

To prove the last assertion on β_∞, by (1.3) it suffices to consider $0 < |\xi_0^{[\varepsilon]}| < \infty$ or $0 < |\widehat{\xi_0^{[\varepsilon]}}| < \infty$, and by the $0-1$-symmetry (Remark 4.5) we need only consider the first case. By monotonicity and translation invariance we can take $\xi_0^{[\varepsilon]} = \delta_0$. The fact that $\mathbf{0}$ is a trap implies $P_{\delta_0}(|\xi_t^{[\varepsilon]}| > 0)$ is non-increasing in t. Therefore (4.11) easily implies $P_{\delta_0}(\tau_{\mathbf{0}} = \infty) > 0$. But $P_{\delta_0}(\tau_{\mathbf{1}} = \infty) = 1$ by (1.3), so we conclude $\beta_\infty(\xi_0^{[\varepsilon]}) > 0$. □

Remark 4.7. The final paragraph in the above proof applies equally well in $d \geq 3$ to show that the conclusion of Theorem 1.2 of [13] may be strengthened to include $\beta_\infty(\xi_0^{[\varepsilon]}) > 0$ if $\xi_0^{[\varepsilon]}$ is not $\mathbf{0}$ or $\mathbf{1}$.

Remark 4.8. Theorems 4.4 and 4.6 hold without the finite range assumptions. That is, we only require that $\{\xi^{[\varepsilon]}, 0 < \varepsilon \leq \varepsilon_0\}$ is a 2-dimensional voter model perturbation in the sense of (1.10)–(1.15) of [13] (where the Hölder rate of convergence in (1.14) of [13] (see (1.14)) is also weakened to (1.15)). Indeed our proofs, and those quoted in [13], only require these conditions. To verify the key condition (4.18), however, we will need to work with finite range voter model perturbations and make a critical assumption on the parameter $\Theta_2 + \Theta_3$ from (3.33).

Theorem 4.9. *Assume $\{\xi^{[\varepsilon]} : 0 < \varepsilon \le \varepsilon_0\}$ is a monotone, asymptotically symmetric finite range voter model perturbation in $\mathbb{Z}^2$ with $\Theta_2 + \Theta_3 > 0$. Then* (4.18) *holds, perhaps with a smaller choice of $\varepsilon_0 > 0$.*

We give the proof in Section 10. It will follow from Theorem 1.15, our weak convergence result to super-Brownian motion with drift $\Theta_2 + \Theta_3$.

We are ready for the proof of our main result, Theorem 1.9, a general complete convergence theorem for monotone cancellative finite range voter model perturbations.

Proof of Theorem 1.9. By Theorem 4.6 it suffices to establish (4.18), perhaps with a smaller $\varepsilon_0 > 0$. By Remark 4.5, $\xi^{[\varepsilon]}$ is symmetric and so in particular is asymptotically symmetric by Remark 3.6. The latter Remark also shows that $r^a = 0$ and therefore $\Theta_2 = 0$. Hence $\Theta_3 + \Theta_2 = \Theta_3 > 0$ (by hypothesis), and so Theorem 4.9 gives (4.18), and we are done. □

To apply Theorem 1.9 it would be useful to have a general, and checkable, sufficient condition for $\Theta_3 > 0$, which would also apply to the q-voter model, and so establish Theorem 1.1 for $d = 2$ as a special case. This is the goal of the next section.

5 Positivity of the drift and the Proof of Theorem 1.1 for $d = 2$

To establish a sufficient condition for $\Theta_3 > 0$ in Theorem 1.9, it will be convenient to first work in a more general setting with any *general* neighbourhood $\mathcal{N}$ (recall from Section 3.1 that $\mathcal{N}$ is finite non-empty subset of $Z^d \setminus \{0\}$) and $d \ge 2$. We consider a strictly subadditive map $r : \{A : A \subset \mathcal{N}\} \to \mathbb{R}$. This means that

$$r(A \cup B) < r(A) + r(B) \text{ for all non-empty disjoint } A, B \subset \mathcal{N}. \tag{5.1}$$

By induction on n this implies that for non empty disjoint sets $A_1, \dots, A_n$ in $\mathcal{N}$,

$$r(\cup_{i=1}^n A_i) \le \sum_{i=1}^n r(A_i), \text{ where strict inequality holds if } n > 1. \tag{5.2}$$

Later we will want to consider a finite range voter model perturbation and take $r = r^s$. Assume for now that $d = 2$. To motivate the above definition recall from (3.42) that for the q-voter model in $d = 2$ we have $r^s(A) = r_{|A|}$, where r_ℓ are as in (1.20). We saw in Section 1.4 (recall (1.24)) that

$$r_{\ell_1+\ell_2} < r_{\ell_1} + r_{\ell_2} \text{ for } 0 < \ell_i, \text{ and } \ell_1 + \ell_2 \le |\mathcal{N}|. \tag{5.3}$$

and so

$$\text{for the 2-dimensional } q\text{-voter model, } r^s \text{ is strictly subadditive.} \tag{5.4}$$

Return now to our earlier setting with $d \ge 2$ and general $\mathcal{N}$. If π is a partition of $\bar{\mathcal{N}}$, $[0]$ denotes the cell of π containing 0 and $|\pi|$ is the cardinality of π. We assume (5.1) and that all sets in a partition are non-empty throughout this section.

Lemma 5.1. *If π is a fixed partition of $\bar{\mathcal{N}}$, then*

$$\sum_{\emptyset \ne A \subset \mathcal{N}} r(A) 1(A \in \pi) \ge \sum_{\emptyset \ne A \subset \mathcal{N}} r(A) 1(\bar{\mathcal{N}} \setminus A = [0]). \tag{5.5}$$

The inequality is strict if $|\pi| > 2$.

Proof. The left-hand side of (5.5) trivially is

$$1(\bar{\mathcal{N}} \setminus [0] \ne \emptyset) \sum_{\emptyset \ne A} r(A) 1(A \in \pi, A \subset \bar{\mathcal{N}} \setminus [0]),$$

while the right-hand side equals

$$r(|\bar{\mathcal{N}} \setminus [0])1(\bar{\mathcal{N}} \setminus [0] \neq \emptyset).$$

So to prove the result we may assume $\bar{\mathcal{N}} \setminus [0] \neq \emptyset$, or equivalently $|\pi| > 1$. Using

$$\bar{\mathcal{N}} \setminus [0] = \cup_{A\in\pi, A\subset\bar{\mathcal{N}}\setminus[0]} A,$$

and the subadditivity (5.2), we have

$$r(\bar{\mathcal{N}} \setminus [0]) \leq \sum_{\emptyset\neq A} r(A)1(A \in \pi, A \subset \bar{\mathcal{N}} \setminus [0]),$$

thus giving (5.5). If $|\pi| = 2$ there is equality in the above, and by (5.2) there is strict inequality if $|\pi| > 2$. □

As an immediate consequence we have:

Lemma 5.2. *If π is a random partition of $\bar{\mathcal{N}}$ such that*

$$P(|\pi| > 2) > 0, \tag{5.6}$$

then

$$E\Big(\sum_{\emptyset\neq A\subset\mathcal{N}} r(A)1(A \in \pi)\Big) > E\Big(\sum_{\emptyset\neq A\subset\mathcal{N}} r(A)1(\bar{\mathcal{N}} \setminus A = [0])\Big). \tag{5.7}$$

Returning to our general r in $d = 2$, and recalling the definitions of $\Theta^{\pm}(A)$ from (1.10), we define $\Theta_3 = \Theta_3(r)$ by

$$\Theta_3 = \sum_{\emptyset\neq A\subset\mathcal{N}} r(A)\Theta^+(A) - \sum_{\emptyset\neq A\subset\mathcal{N}} r(A)\Theta^-(A) := \Theta_3^+ - \Theta_3^-. \tag{5.8}$$

Note that this agrees with our earlier definition of Θ_3 in (3.33) if $r = r^s$ for the finite range voter perturbations in Section 3.2. To use the above to show the positivity of Θ_3 in (5.8), recall $K_3(A_1, A_2, A_3)$ from Proposition 1.6 and the notation $\mathcal{P}_k(\Gamma)$ and $\mathcal{P}(\Gamma)$ from Section 1.3. Recall also that (5.1) is still in force and the q-voter drift Θ in (1.22) corresponds to the special case $r(A) = r_{|A|}$ with r_ℓ as in (1.20).

Corollary 5.3. *If $d = 2$, then $\Theta_3 > 0$, and, in particular, Θ in (1.22) is also strictly positive.*

Proof. Let $\kappa = \sum_{\{A_1,A_2,A_3\}\in\mathcal{P}_3(\bar{\mathcal{N}})} K_3(A_1, A_2, A_3) > 0$, where the sum is over sets, not ordered triples, and the positivity is clear by $|\bar{\mathcal{N}}| \geq 5$ (see (1.2)). Define a random partition in $\mathcal{P}_3(\bar{\mathcal{N}})$ by

$$P(\pi = \{A_1, A_2, A_3\}) = K_3(A_1, A_2, A_3)/\kappa.$$

Both are well-defined by the symmetry of K_3, and (5.6) holds because $|\pi| = 3$ a.s. If A is a non-empty subset of $\bar{\mathcal{N}}$, then

$$P(A \in \pi) = \sum_{\{A_1,A_2\}\in\mathcal{P}(\bar{\mathcal{N}}\setminus A)} P(\pi = \{A, A_1, A_2\}) = \sum_{\{A_1,A_2\}\in\mathcal{P}(\bar{\mathcal{N}}\setminus A)} K_3(A, A_1, A_2)/\kappa. \tag{5.9}$$

Therefore from (5.8)

$$\Theta_3^+ := \sum_{\emptyset\neq A\subset\mathcal{N}} r(A)\Theta^+(A) = \kappa \sum_{\emptyset\neq A\subset\mathcal{N}} r(A)P(A \in \pi) = \kappa E\Big(\sum_{\emptyset\neq A\subset\mathcal{N}} r(A)1(A \in \pi)\Big). \tag{5.10}$$

Similarly, apply (5.9) with $\bar{\mathcal{N}} \setminus A$ in place of A to see that

$$\Theta_3^- := \sum_{\emptyset \neq A \subset \mathcal{N}} r(A)\Theta^-(A) = \kappa \sum_{\emptyset \neq A \subset \mathcal{N}} r(A) P(\bar{\mathcal{N}} \setminus A \in \pi) = \kappa E\Big(\sum_{\emptyset \neq A \subset \mathcal{N}} r(A) 1(\bar{\mathcal{N}} \setminus A \in \pi)\Big).$$

Note that for $A \subset \mathcal{N}$ we have $0 \in \bar{\mathcal{N}} \setminus A$, and so $\bar{\mathcal{N}} \setminus A \in \pi$ iff $[0] = \bar{\mathcal{N}} \setminus A$. This shows that the above implies

$$\Theta_3^- = \kappa E\Big(\sum_{\emptyset \neq A \subset \mathcal{N}} r(A) 1(\bar{\mathcal{N}} \setminus A = [0])\Big). \tag{5.11}$$

So by Lemma 5.2, (5.10) and (5.11) we have $\Theta_3^+ > \Theta_3^-$, and therefore, $\Theta_3 > 0$. As noted above, the positivity of Θ in (1.22) follows from the special case $r(A) = r_{|A|}$. □

Remark 5.4. If $r(A) = |A|$ (not strictly subadditive!) one easily sees that equality holds in Lemmas 5.1 and 5.2, the latter without any condition on π. The above proof then shows that

$$\sum_{\emptyset \neq A \subset \mathcal{N}} |A|(\Theta^+(A) - \Theta^-(A)) = 0. \tag{5.12}$$

This identity simplified some of the formulae for Θ_3 in the examples of Section 3.2.

If $r(A) = 1$ for all $A \subset \mathcal{N}$ (which is strictly subadditive), then from (5.10) we have $\Theta_3^+ = 2\kappa$ because there are exactly two subsets of $\mathcal{N}$ in π corresponding to the two sets in π other than $[0]$. Similarly from (5.11) we get that $\Theta_3^- = \kappa$ because there is exactly one subset of $\mathcal{N}$ whose complement in $\bar{\mathcal{N}}$ is $[0]$, namely $\bar{\mathcal{N}} \setminus [0]$. Therefore

$$\sum_{\emptyset \neq A \subset \mathcal{N}} (\Theta^+(A) - \Theta^-(A)) = \kappa = \sum_{\{A_1,A_2,A_3\} \in \mathcal{P}_3(\bar{\mathcal{N}})} K_3(A_1,A_2,A_3) = \lim_{t\to\infty} (\log t)^3 \hat{P}(|B_t^{\bar{\mathcal{N}}}| = 3) > 0. \tag{5.13}$$

The last equality follows easily from the definition of K_3 by decomposing $\hat{P}(|B_t^{\bar{\mathcal{N}}}| = 3)$ into the possible partitions induced by the sets of sites which have coalesced at time t and taking limits.

The following corollary is immediate from Theorem 1.9 and Corollary 5.3. It represents our simplest criteria for a CCT to hold in $d = 2$.

Corollary 5.5. *Assume $d = 2$, $\{\xi^{[\varepsilon]} : 0 < \varepsilon \le \varepsilon_0\}$ satisfies (4.14) and r^s, given by (1.18), is strictly subadditive (i.e., (5.1) holds for $r = r^s$). Then $\Theta_3 > 0$ and there is an $\varepsilon_1 > 0$ such that for $\varepsilon \in (0, \varepsilon_1)$, the complete convergence theorem with coexistence (CCT) holds for $\xi^{[\varepsilon]}$.*

To illustrate the use of the Corollary, we first show how it quickly gives Theorem 1.1 (which was already outlined in Section 1.4).

Proof of Theorem 1.1 for $d = 2$. We apply Corollary 5.5 above with $\xi^{[\varepsilon]}$ the $(1-\varepsilon)$-voter model, $\xi^{(1-\varepsilon)}$. The cancellative property for $|\mathcal{N}| \le 8$ is shown in Lemma 3.3 for ε small enough, and the finite range voter perturbation property is established in Example 3.8. The monotonicity of any q-voter model is elementary (recall (1.4)). Strict subadditivity of r^s was already noted in (5.4) and so Corollary 5.5 gives the result. □

Recall from the examples at the end of Section 3.2 that for the Lotka-Volterra models, affine voter models and geometric voter models, we have for non-empty $A \subset \mathcal{N}$, $r^s(A) = -(|A|/|\mathcal{N}|)^2$, $r^s(A) = -\frac{|A|}{|\mathcal{N}|} + 1$, and $r^s(A) = \frac{|A|(|\mathcal{N}|-|A|)}{2|\mathcal{N}|}$, respectively. All of these asymptotic rate functions are strictly subadditive, as one can easily check. Condition (4.14) was verified for all of these models in Examples 3.9-3.11, where for the Lotka-Volterra model we take $\alpha_1 = \alpha_2 \ge 1/2$. The following theorems are then also immediate consequences of Corollary 5.5. The first is the $d = 2$ case of Theorem 1.1 of [13] which helped motivate the general result here.

Theorem 5.6. *Let $d = 2$, let $\mathcal{N}$ be a neighbourhood, and let LV(α) denote the Lotka-Volterra model with parameters $\alpha_1 = \alpha_2 = \alpha$. Then $\Theta_3^{lv} > 0$ and there is an $\alpha_c \in (0,1)$ such that for all $\alpha \in (\alpha_c, 1)$, the complete convergence theorem with coexistence holds for LV(α).*

Theorem 5.7. *Let $d = 2$, let $\mathcal{N}$ be a neighbourhood, and let AV(α) denote the affine voter model with parameter α. Then $\Theta_3^{av} > 0$ and there is an $\alpha_c \in (0,1)$ such that for all $\alpha \in (\alpha_c, 1)$, the complete convergence theorem with coexistence holds for AV(α).*

Theorem 5.8. *Let $d = 2$, let $\mathcal{N}$ be a neighbourhood, and let GV(θ) denote the geometric voter model with parameter θ. Then $\Theta_3^{gv} > 0$ and there is a $\theta_c \in (0,1)$ such that for all $\theta \in (\theta_c, 1)$, the complete convergence theorem with coexistence holds for GV(θ).*

Finally, we give the promised direct proof of $f'(0) > 0$ for the q voter model and $d \geq 3$. Recall that in this case r_ℓ is as in (1.20),

$$c^*(x,\xi) = \sum_{\ell=1}^{|\mathcal{N}|} r_l\Big(\widehat{\xi}(x)1\{n_1(x,\xi) = \ell\} + \xi(x)1\{n_0(x,\xi) = \ell\}\Big), \tag{5.14}$$

and f is given by (3.51).

Proposition 5.9. *([4]) Assume $d \geq 3$ and $\mathcal{N}$ is a fixed neighbourhood. Then $f'(0) > 0$.*

Proof. It follows from (3.51) and (5.14) that

$$\begin{aligned} f(u) &= \Big\langle \sum_{\ell=1}^{|\mathcal{N}|} r_\ell \widehat{\xi}(0)1(n_1(0,\xi) = \ell)\Big\rangle_u - \Big\langle \sum_{\ell=1}^{|\mathcal{N}|} r_\ell \xi(0)1(n_0(0,\xi) = \ell)\Big\rangle_u \\ &:= f_1(u) - f_0(u). \end{aligned} \tag{5.15}$$

Let $\{B^x : x \in \bar{\mathcal{N}}\}$ be the system of coalescing random walks introduced in Section 1.3 under $\hat{P}$, but now in dimension $d \geq 3$, and let $\{\xi_0^u(x) : x \in \mathbb{Z}^d\}$ have Bernoulli product measure with density $u \in [0,1]$. Let $\pi_t \in \mathcal{P}(\bar{\mathcal{N}})$ be the random partition determined by the coalescing random walks $\{B_t^x, x \in \bar{\mathcal{N}}\}$ using the equivalence relation $x \sim_t y$ iff $\sigma(x,y) = \inf\{u : B_s^x = B_u^y\} \leq t$, and let $\pi_\infty = \lim_{t\to\infty} \pi_t$. In this way $x \sim_\infty y$ iff $\sigma(x,y) < \infty$ is the associated equivalence relation. If $A \subset \bar{\mathcal{N}}$ and $T \in [0,\infty]$, let

$$[A]_T = \{\lambda \in \pi_T : \lambda \cap A \neq \emptyset\},$$

and, abusing this notation slightly, write $[x]_T$ for the cell of π_T containing x. Use the duality between the voter model and coalescing random walk to see that

$$\begin{aligned} f_1(u) &= \Big\langle \sum_{\emptyset \neq A \subset \mathcal{N}} r_{|A|} 1(\xi|_A = 1, \xi|_{\bar{\mathcal{N}}\setminus A} = 0)\Big\rangle_u \\ &= \lim_{T\to\infty} \sum_{\emptyset \neq A \subset \mathcal{N}} r_{|A|} \sum_{i=1}^{|A|} \sum_{j=1}^{|\mathcal{N}|+1-|A|} \hat{P}(B_T^A \subset \{x : \xi_0^u(x) = 1\}, |B_T^A| = i, \\ &\qquad\qquad B_T^{\bar{\mathcal{N}}\setminus A} \subset \{x : \xi_0^u(x) = 0\}, |B_T^{\mathcal{N}\setminus A}| = j) \\ &= \sum_{\emptyset \neq A \subset \mathcal{N}} r_{|A|} \sum_{i=1}^{|A|} \sum_{j=1}^{|\mathcal{N}|+1-|A|} \lim_{T\to\infty} \hat{P}(\sigma(A, \bar{\mathcal{N}} \setminus A) > T,\ |[A]_T| = i,\ |[\bar{\mathcal{N}} \setminus A]_T| = j) u^i (1-u)^j \\ &= \sum_{\emptyset \neq A \subset \mathcal{N}} r_{|A|} \sum_{i=1}^{|A|} \sum_{j=1}^{|\mathcal{N}|+1-|A|} \hat{P}(\sigma(A, \bar{\mathcal{N}} \setminus A) = \infty,\ |[A]_\infty| = i,\ |[\bar{\mathcal{N}} \setminus A]_\infty| = j) u^i (1-u)^j. \end{aligned} \tag{5.16}$$

Differentiate the above at $u=0$ and so conclude that

$$f_1'(0) = \sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}\hat{P}(\sigma(A,\bar{\mathcal{N}}\setminus A)=\infty,\ |[A]_\infty|=1).$$

Note that $\sigma(A,\bar{\mathcal{N}}\setminus A)=\infty$ iff A is a union of cells in π_∞ and so

$$(\sigma(A,\bar{\mathcal{N}}\setminus A)=\infty \text{ and } |[A]_\infty|=1) \iff A\in\pi_\infty.$$

The above expression for $f_1'(0)$ now becomes

$$f_1'(0) = \hat{E}\Big(\sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}1(A\in\pi_\infty)\Big). \tag{5.17}$$

In a similar way to (5.16), we get

$$\begin{aligned} f_0(u) &= \Big\langle \sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}1(\xi|_A=0,\xi_{\bar{\mathcal{N}}\setminus A}=1)\Big\rangle_u \\ &= \sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}\sum_{i=1}^{|A|}\sum_{j=1}^{|\mathcal{N}|+1-|A|} \hat{P}(\sigma(A,\bar{\mathcal{N}}\setminus A)=\infty,\ |[A]_\infty|=i,\ |[\bar{\mathcal{N}}\setminus A]_\infty|=j)(1-u)^iu^j. \end{aligned}$$

Differentiating at $u=0$ we get

$$\begin{aligned} f_0'(0) &= \sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}\hat{P}(\sigma(A,\bar{\mathcal{N}}\setminus A)=\infty,\ |[\bar{\mathcal{N}}\setminus A]_\infty|=1) \\ &= \hat{E}\Big(\sum_{\emptyset\neq A\subset\mathcal{N}} r_{|A|}1(\bar{\mathcal{N}}\setminus A=[0]_\infty)\Big). \end{aligned} \tag{5.18}$$

For the last, note that for $A\subset\mathcal{N}$, the event inside the $\hat{P}$ is precisely $\{\bar{\mathcal{N}}\setminus A=[0]_\infty\}$.

Note that $\hat{P}(|\pi_\infty|=|\mathcal{N}|+1)>0$ for $d\geq 3$ as there is positive probability none of the walks coalesce, and so (5.6) holds. We now may apply Lemma 5.2 to conclude from (5.17) and (5.18) that $f_1'(0)>f_0'(0)$, and so $f'(0)>0$. □

Remark 5.10. The above proof applies equally well to any $\{r_\ell : 1\leq\ell\leq|\mathcal{N}|\}$ satisfying (5.3). In fact, with only notational changes, it applies to r as in (5.1) where f is given by (3.51) and

$$c^*(x,\xi) = \widehat{\xi}(0)\sum_{\emptyset\neq A\subset\mathcal{N}} r(A)1(\xi|_{\mathcal{N}}=1_A) + \xi(0)\sum_{\emptyset\neq A\subset\mathcal{N}} r(A)1(\xi|_{\mathcal{N}}=1_{\mathcal{N}\setminus A}). \tag{5.19}$$

6 A general convergence theorem to super-Brownian motion in two dimensions

In Section 4 we extended the SDE construction of our particle systems to allow for infinite initial conditions and also simultaneously deal with $\widehat{\xi}^{[\varepsilon]}$. In this setting we have no need to deal with these extensions and so no longer require monotonicity or have to deal with the equation (4.7) for $\widehat{\xi}$. Assume the walk kernel, p, is as in the beginning of Section 1.4 and $\{\xi^{[\varepsilon]} : 0<\varepsilon\leq\varepsilon_0\}$ is an asymptotically symmetric finite range voter model perturbation as in Section 3.2. This assumption will be in force throughout the rest of this work. We continue to use notation from Section 3.2 and, as in that Section, $N\geq N(\varepsilon_0)>e^3$ is our fundamental parameter where $\varepsilon=\varepsilon_N=(\log N)^3/N$. Recall from (3.23) that the rescaled voter model perturbation, ξ^N in (3.19) has rate function

$$c^N(x,\xi^{(N)}) = Nc^{N,\mathsf{vm}}(x,\xi^{(N)}) + (\log N)c^{N,a}(x,\xi^{(N)}) + (\log N)^3c^{N,s}(x,\xi^{(N)}), \qquad x\in S_N,\ \xi^{(N)}\in\{0,1\}^{S_N}, \tag{6.1}$$

where $c^{N,s}$ and $c^{N,a}$ are as in (3.20) and (3.21), respectively. Assume throughout that $|\xi_0^N| < \infty$.

By Remark 4.1 we can construct ξ^N, as in Section 4, as the unique $(\mathcal{F}_t^N)$-adapted solution of

$$\begin{aligned}\xi_t^N(x) = \xi_0^N(x) &+ \int_0^t \int (1-\xi_{s-}^N(x))1(u \le c^N(x,\xi_{s-}^N))N^{x,0}(ds,du) \\ &- \int_0^t \int \xi_{s-}^N(x)1(u \le c^N(x,\xi_{s-}^N))N^{x,1}(ds,du), \quad t \ge 0,\ x \in S_N. \end{aligned} \tag{6.2}$$

Then ξ^N is the unique Feller process associated with the rate function $c^N(x,\xi)$. Here $\{N^{x,i}, x \in S_N, i = 0,1\}$ are independent Poisson point processes on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $ds \times du$ and $\{\mathcal{F}_t^N, t \ge 0\}$ is the natural right-continuous filtration generated by these point processes.

We again use this setting to couple spin-flip systems but now in a different manner from Section 4. If $\bar{c}^N(x,\xi)$ is a rate function, also satisfying (4.2) and (4.3), such that for $\xi \le \bar{\xi}$,

$$\begin{aligned}&\bar{c}^N(x,\bar{\xi}) \ge c^N(x,\xi) \text{ if } \xi(x) = \bar{\xi}(x) = 0,\\ &\bar{c}^N(x,\bar{\xi}) \le c^N(x,\xi) \text{ if } \xi(x) = \bar{\xi}(x) = 1,\end{aligned} \tag{6.3}$$

and $\bar{\xi}_t^N$ is constructed as in (6.2) using $\bar{c}^N(x,\xi)$, then

$$\bar{\xi}_0^N \ge \xi_0^N \text{ implies } \bar{\xi}_t^N \ge \xi_t^N \text{ for all } t \ge 0. \tag{6.4}$$

In Proposition 2.1 of [11] this is proved under monotonicity of ξ^N when ξ^N is a killed version of $\bar{\xi}^N$, but the same elementary argument applies without monotonicity under (6.3).

As in (1.32), define the measure-valued process associated with $\xi_t^N(x),\ x \in S_N$ by

$$X_t^N = (1/N') \sum_{x \in S_N} \xi_t^N(x)\delta_x. \tag{6.5}$$

We use the SDE (6.2) to see that X^N satisfies a martingale problem reminiscent of that of SBM. Introduce the scaled probability kernel

$$p_N(x) = p(x\sqrt{N}), \quad x \in S_N. \tag{6.6}$$

Unless otherwise indicated, assume $\Phi \in C_b([0,T]\times\mathbb{R}^2)$ is such that $\dot{\Phi} := \frac{\partial \Phi}{\partial t} \in C_b([0,T]\times\mathbb{R}^2)$. Define

$$A_N\Phi(s,x) = \sum_{y \in S_N} Np_N(y-x)\left(\Phi(s,y) - \Phi(s,x)\right),$$

and

$$D_t^{N,1}(\Phi) = \int_0^t X_s^N\big(A_N\Phi(s,\cdot) + \dot{\Phi}(s,\cdot)\big)ds.$$

Introduce

$$\ell_N^{(j)} = \begin{cases}\log N & \text{if } j = 2,\\ (\log N)^3 & \text{if } j = 3.\end{cases}$$

To be consistent with the notation in [9] for $x \in S_N$ and $\xi \in \{0,1\}^{S_N}$ we define

$$\begin{aligned}d^{N,2}(x,\xi) &= \widehat{\xi}(x)c^{N,a}(x,\xi)\\ d^{N,3}(x,\xi) &= \widehat{\xi}(x)c^{N,s}(x,\xi) - \xi(x)c^{N,s}(x,\xi)\end{aligned} \tag{6.7}$$

and for $j = 2, 3$,

$$d^{N,j}(s, \xi, \Phi) = \frac{\ell_N^{(j)}}{N'} \sum_{x \in S_N} \Phi(s, x) d^{N,j}(x, \xi), \tag{6.8}$$

$$D_t^{N,j}(\Phi) = \int_0^t d^{N,j}(s, \xi_s^N, \Phi) ds. \tag{6.9}$$

One may then use the stochastic calculus for Poisson integrals and integration by parts to rewrite $\xi_t^N(x)\Phi(t, x)$, just as in Propositions 2.2 and 2.3 in [10], to see that

$$X_t^N(\Phi(t, \cdot)) = X_0^N(\Phi(0, \cdot)) + D_t^{N,1}(\Phi) + D_t^{N,2}(\Phi) + D_t^{N,3}(\Phi) + M_t^N(\Phi), \tag{6.10}$$

where $M_t^N(\Phi)$ is a square integrable $(\mathcal{F}_t^N)$-martingale with previsible square function

$$\langle M^N(\Phi)\rangle_t = \langle M^N(\Phi)\rangle_{1,t} + \langle M^N(\Phi)\rangle_{2,t}, \tag{6.11}$$

with

$$\begin{aligned} \langle M^N(\Phi)\rangle_{1,t} &= \int_0^t \frac{\log N}{N'} \sum_{x \in S_N} \Phi(s, x)^2 c^{N,\mathsf{vm}}(x, \xi_s^N) ds, \\ \langle M^N(\Phi)\rangle_{2,t} &= \int_0^t \frac{1}{N'^2} \sum_{x \in S_N} \Phi(s.x)^2 [\ell_N^{(2)} c^{N,a}(x, \xi_s^N) + \ell_N^{(3)} c^{N,s}(x, \xi_s^N)] \, ds. \end{aligned} \tag{6.12}$$

Note that in spite of the suggestive notation the last term may in fact be negative. The above three displays are reminiscent of the martingale problem (MP) for a super-Brownian motion in Section 1.5. In the next two sections we will take term by term limits in (6.10) to establish Theorem 1.15, our general weak convergence result. We restate it below for convenience. Recall that σ^2 is as in (3.13) and Θ_2, Θ_3 are defined in (3.33).

Theorem 6.1. *Assume $\{\xi^{[\varepsilon]} : 0 < \varepsilon \le \varepsilon_0\}$ is an asymptotically symmetric finite range voter model perturbation on $\mathbb{Z}^2$. If $X_0^N \to X_0$ in $\mathcal{M}_F$, then*

$$X^N \Rightarrow \textit{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \Theta_2 + \Theta_3) \text{ in the Skorokhod space } D(\mathbb{R}_+, \mathcal{M}_F) \text{ as } N \to \infty.$$

In Remark 1.16 we showed how Theorem 1.11, the rescaled limit theorem for 2-dimensional q-voter models, follows from the above. We now describe a number of other corollaries, all of course in two dimensions.

Example 6.2. (Lotka-Volterra Models). Recall from Example 3.9 that the Lotka-Volterra Models discussed there constituted an asymptotically symmetric finite range voter model. The kernel p is $1_{\mathcal{N}}/|\mathcal{N}|$ and so σ^2 is as in (1.1). Recall also that $\xi^{[\varepsilon_N]}$ denotes a Lotka-Volterra model with parameters (α_1^N, α_2^N), where

$$\alpha_i^N = 1 - \varepsilon_N + \beta_i^{(\varepsilon_N)}(\log(1/\varepsilon_N))^{-2}\varepsilon_N = 1 - \frac{(\log N)^3}{N} + \beta_i^N \frac{\log N}{N},$$

and $\beta_i^N \to \beta_i \in \mathbb{R}$ as $N \to \infty$. Let

$$\xi_t^N(x) = \xi_{Nt}^{[\varepsilon_N]}(x\sqrt{N}) \text{ for } x \in S_N, \text{ and } X_t^N = \frac{1}{N'} \sum_{x \in S_N} \xi_t^N(x)\delta_x. \tag{6.13}$$

If $X_0^N \to X_0 \in \mathcal{M}_F(\mathbb{R}^2)$ as $N \to \infty$, then Theorem 6.1 implies that

$$X^N \Rightarrow \mathrm{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \Theta_2^{\mathsf{lv}} + \Theta_3^{\mathsf{lv}}),$$

where Θ_i^{lv} are as in Example 3.9. We can write these drifts in another way. Let e_1, e_2 be iid rv's uniformly distributed over $\mathcal{N}$ and, using the notation from Section 1.3, set

$$K = E(K_3(0, e_1, e_2)1(e_1 \neq e_2)), \text{ and } \gamma = E(K_2(\{e_1, e_2\}, \{0\})). \tag{6.14}$$

Then with a bit of work one can show that $\Theta_3^{\mathrm{lv}} = K$ and $\Theta_2^{\mathrm{lv}} = (\beta_0 - \beta_1)\gamma$. For example, in the proof of the latter, the summand A arising in Θ_2^{lv} will be $[e_1]_t$ (recall from the proof of Proposition 5.9 this is the set of initial conditions in $\mathcal{N}$ that have coalesced with e_1 by time t) and then let $t \to \infty$. In this way the drift of the limiting SBM becomes $K + (\beta_0 - \beta_1)\gamma$. This is the form of the limit derived in Theorem 1.5 of [9] whose proof will play an important role in the derivation of Theorem 6.1 to come.

Example 6.3. (Affine Voter Models). In Example 3.10 we showed the affine voter models are an asymptotically symmetric voter model perturbation with kernel $p(x) = 1_{\mathcal{N}}(x)/|\mathcal{N}|$, σ^2 as in (1.1), $\Theta_2 = 0$, and $\Theta_3^{\mathrm{av}} = \kappa = \lim_{t\to\infty}(\log t)^3 \hat{P}(|B_t^{\mathcal{N}}| = 3) > 0$ (recall (5.13)). Let $\xi_{Nt}^{[\varepsilon_N]}(x)$, $x \in \mathbb{Z}^2$ be an affine voter model with parameter $\alpha = 1 - \varepsilon_N$, and define ξ^N and X^N as in (6.13). Theorem 6.1 implies that

$$\text{if } X_0^N \to X_0 \in \mathcal{M}_F(\mathbb{R}^2), \text{ then } X^N \Rightarrow \mathrm{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \kappa).$$

Example 6.4. (Geometric Voter Models). By Example 3.11, the geometric voter models give an asymptotic symmetric voter model perturbation with p and σ^2 as above, $\Theta_2 = 0$ and $\Theta_3^{\mathrm{gv}} = \frac{|\mathcal{N}|}{2}\Theta_3^{\mathrm{lv}} = \frac{|\mathcal{N}|}{2}K$ where $K > 0$ is as in (6.14). Let $\xi^{[\varepsilon_N]}$ be a geometric voter model with parameter $\theta = 1 - \varepsilon_N$ and assume ξ^N and X^N are as in (6.13). Theorem 6.1 implies that

$$\text{if } X_0^N \to X_0 \in \mathcal{M}_F(\mathbb{R}^2), \text{ then } X^N \Rightarrow \mathrm{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \frac{|\mathcal{N}|}{2}K).$$

Remark 6.5. We have started with a finite range voter model perturbation $\{c_\varepsilon : 0 < \varepsilon \leq \varepsilon_0\}$ and rescaled to obtain rates c^N as in (6.1). At times it may be more natural to start with the rescaled rates c^N for $N \geq N_0 \geq e^3$. Assume now that

$$c^N(x, \xi^{(N)}) = Nc^{N,\mathrm{vm}}(x, \xi^{(N)}) + (\log N)c^{N,a,*}(x, \xi^{(N)}) + (\log N)^3 c^{N,s,*}(x, \xi^{(N)}),$$
$$x \in S_N,\ \xi^{(N)} \in \{0,1\}^{S_N},$$

where for some $\mathbb{R}$-valued functions $g_i^{N,a,*}$, $g^{N,s,*}$ on $\{0,1\}^{\mathcal{N}}$, $i = 0, 1$,

$$c^{N,a,*}(x, \xi^{(N)}) = \widehat{\xi}(x\sqrt{N})g_1^{N,a,*}(\xi|_{x\sqrt{N}+\mathcal{N}}) + \xi(x\sqrt{N})g_0^{N,a,*}(\xi|_{x\sqrt{N}+\mathcal{N}}), \tag{6.15}$$

and

$$c^{N,s,*}(x, \xi^{(N)}) = \widehat{\xi}(x\sqrt{N})g^{N,s,*}(\widehat{\xi}|_{x\sqrt{N}+\mathcal{N}}) + \xi(x\sqrt{N})g^{N,s,*}(\xi|_{x\sqrt{N}+\mathcal{N}}). \tag{6.16}$$

We assume there are functions $g_i^{a,*}$, $g^{s,*}$ on $\{0,1\}^{\mathcal{N}}$, $i = 0, 1$, such that

$$\lim_{N\to\infty}\Big(\sum_{i=0}^{1} \|g_i^{a,N,*} - g_i^{a,*}\|_\infty\Big) + \|g^{s,N,*} - g^{s,*}\|_\infty = 0,$$

and also that $\mathbf{0}$ and $\mathbf{1}$ are traps, that is, for all N and $x \in S_N$, $c^N(x, \mathbf{0}) = c^N(x, \mathbf{1}) = 0$. This setting clearly includes the c^N arising in (6.1), and in fact appears to be more general since (6.1) requires $\xi(x\sqrt{N})c^{N,a}(x, \xi^{(N)}) = 0$. To see that it is in fact included in (6.1), define for $\xi \in \{0,1\}^{\mathcal{N}}$,

$$g_0^{\varepsilon_N}(\xi) = g^{N,s,*}(\xi) + (\log N)^{-2} g_0^{N,a,*}(\xi) \to g^{s,*}(\xi) \text{ as } N \to \infty, \tag{6.17}$$

and

$$g_1^{\varepsilon_N}(\xi) = g_0^{\varepsilon_N}(\widehat{\xi}) + (\log N)^{-2}[g_1^{N,a,*}(\xi) - g_0^{N,a,*}(\widehat{\xi})] \to g^{s,*}(\widehat{\xi}) \text{ as } N \to \infty. \tag{6.18}$$

Then one easily checks that c^N is as in (6.1) where $c^{N,s}$ and $c^{N,a}$ are given by (3.20) and (3.21), respectively, in terms of the $g_i^{\varepsilon_N}$ given above. Moreover,

$$\begin{aligned} g^a(\xi) &:= \lim_{N\to\infty} (\log(1/\varepsilon_N))^2 (g_1^{\varepsilon_N}(\xi) - g_0^{\varepsilon_N}(\widehat{\xi})) \\ &= \lim_{N\to\infty} (\log(1/\varepsilon_N))^2 (\log N)^{-2} [g_1^{N,a,*}(\xi) - g_0^{N,a,*}(\widehat{\xi})] = g_1^{a,*}(\xi) - g_0^{a,*}(\widehat{\xi}). \end{aligned}$$

Use c^N to define c_{ε_N} as in (3.18). Then $\{c_\varepsilon : 0 < \varepsilon \le \varepsilon_0 := \varepsilon_{N_0}\}$ is an asymptotically symmetric voter model perturbation. Recalling (3.28), for $A \subset \mathcal{N}$ we have

$$r^a(A) = g^a(1_A) = g_1^{a,*}(1_A) - g_0^{a,*}(1_{\mathcal{N}\setminus A}), \tag{6.19}$$

while (6.17) and (3.30) give

$$r^s(A) = \lim_{N\to\infty} g_0^{\varepsilon_N}(1_{\mathcal{N}\setminus A}) = g^{s,*}(1_{\mathcal{N}\setminus A}). \tag{6.20}$$

If $\xi^{[\varepsilon]}$ is the process with rate c_ε, then (3.18) implies that $\xi_t^N(x) = \xi_{N\cdot}^{[\varepsilon_N]}(\sqrt{N}\cdot)$ ($x \in S_N$) has rate function c^N. Therefore if X^N is as in (6.5), we may apply Theorem 6.1 to conclude that
$X^N \Rightarrow \mathrm{SBM}(X_0, 4\pi\sigma^2, \sigma^2, \Theta_2 + \Theta_3)$, where (by (6.19) and (6.20))

$$\Theta_2 = \sum_{\emptyset\neq A\subset\mathcal{N}} (g_1^{a,*}(1_A) - g_0^{a,*}(1_{\mathcal{N}\setminus A}))K_2(A, \bar{\mathcal{N}}\setminus A),\ \Theta_3 = \sum_{\emptyset\neq A\subset\mathcal{N}} g^{s,*}(1_{\mathcal{N}\setminus A})(\Theta^+(A) - \Theta^-(A)). \tag{6.21}$$

Example 6.6. As a specific example of the above, recall the lower order weak limit theorem for two-dimensional Lotka-Volterra models in [12], where now $\alpha_i^N = 1 + \frac{\log N}{N}\beta_i^N$ and $\beta_i^N \to \beta_i$, $i = 0, 1$. We continue to assume $p(x) = 1_{\mathcal{N}}/|\mathcal{N}|$. This led to rescaled rates (see (1.6) in [12])

$$c^N(x, \xi^{(N)}) = Nc^{N,\mathtt{vm}}(x, \xi^{(N)}) + \log N\Big[\widehat{\xi}(x\sqrt{N})\beta_0^N f_1(x\sqrt{N}, \xi)^2 + \xi(x\sqrt{N})\beta_1^N f_0(x\sqrt{N}, \xi)^2\Big].$$

So comparing with (6.15) and (6.16) above we have $c^{N,s,*} = g^{x,*} = 0$, and $g_1^{a,*}(\xi) = \beta_0 f_1(0,\xi)^2$ and $g_0^{a,*}(\widehat{\xi}) = \beta_1 f_0(0,\widehat{\xi})^2 = \beta_1 f_1(x,\xi)^2$. So (6.21) implies $\Theta_3 = 0$ and

$$\Theta_2 = (\beta_0 - \beta_1) \sum_{\emptyset\neq A\subset\mathcal{N}} \Big(\frac{|A|}{|\mathcal{N}|}\Big)^2 K_2(A, \bar{\mathcal{N}} \setminus A) = \Theta_2^{\mathtt{lv}} = (\beta_0 - \beta_1)\gamma,$$

where γ is as in Example 6.2, which also gives the above expression for $\Theta_2^{\mathtt{lv}}$. Therefore by Remark 6.5 and Theorem 6.1, if we define ξ^N, and X^N as in Example 6.2 and $X_0^N \to X_0$, then

$$X^N \Rightarrow \mathrm{SBM}(X_0, 4\pi\sigma^2, \sigma^2, (\beta_0 - \beta_1)\gamma).$$

This is Theorem 1.2 of [12] but with a seemingly different parameter, γ, in place of the γ^* in [12]. It is, however, easy to use Proposition 2.2 of [12] and (1.9) with $n = 2$ to check that $\gamma^* = \gamma$.

7 Controlling the drift terms, and total mass bounds

The goal in this section (Proposition 7.16) is to show that for $j = 2, 3$, the drift terms $D_t^{N,j}(\Phi)$ arising in (6.10) behave asymptotically like $\Theta_j \int_0^t X_s^N(\Phi_s)\,ds$, where $\Phi_s(x) = \Phi(s, x)$. Use notation from the previous section. In particular, ξ^N is as in (3.19) with rate function, c^N, as in (6.1).

7.1 Small time comparison bounds

We begin with some elementary bounds on the drifts. Let $\mathcal{N}_N = \mathcal{N}/\sqrt{N}$. By the definitions of $d^{N,j}(x,\xi)$, (3.27) and (3.29), we have

$$\begin{aligned} d^{N,2}(x,\xi) &= \widehat{\xi}(x) \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,a}(\sqrt{N}A) 1\{\xi|_{x+\mathcal{N}_N} = 1_{x+A}\}, \\ &= \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,a}(\sqrt{N}A) 1\{\xi|_{x+\bar{\mathcal{N}}_N} = 1_{x+A}\}, \\ d^{N,3}(x,\xi) &= \widehat{\xi}(x) \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A) 1\{\xi|_{x+\mathcal{N}_N} = 1_{x+A}\} - \xi(x) \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A) 1\{\xi|_{x+\mathcal{N}_N} = 1_{x+\mathcal{N}_N \setminus A}\} \\ &= \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A) 1\{\xi|_{x+\bar{\mathcal{N}}_N} = 1_{x+A}\} - \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A) 1\{\xi|_{x+\bar{\mathcal{N}}_N} = 1_{x+\bar{\mathcal{N}}_N \setminus A}\}. \end{aligned} \tag{7.1}$$

Observe that at most one term in each of the above sums can be nonzero, and this would require that $\xi(y) = 1$ for some $y \in x + \bar{\mathcal{N}}_N$. This implies that for $j = 2, 3$,

$$|d^{N,j}(x,\xi)| \le \|r\| 1\{\xi(y) = 1 \text{ for some } y \in x + \bar{\mathcal{N}}_N\} \le \|r\| \sum_{y \in x + \bar{\mathcal{N}}_N} \xi(y), \tag{7.2}$$

and hence

$$|d^{N,j}(x,\xi)| \le \|r\|, \tag{7.3}$$

and for $\Phi : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$,

$$|d^{N,j}(s,\xi,\Phi)| \le \|r\| |\Phi_s|_\infty |\bar{\mathcal{N}}| \frac{\ell_N^{(j)}}{N'} \sum_{x \in S_N} \xi(x). \tag{7.4}$$

To control the drifts over small time intervals we will condition $d^{N,2}(s,\xi_s^N,\Phi)$ and $d^{N,3}(s,\xi_s^N,\Phi)$ on $\mathcal{F}_{s-u_N}^N$ for small u_N and compare this small time evolution of ξ^N with a voter model, for which we can make explicit calculations using duality.

Set $f_i^{(N)}(x,\xi) = \sum_{y \in S_N} p_N(y-x) 1\{\xi(y) = i\}$, $n_i^{(N)}(x,\xi) = \sum_{y \in \mathcal{N}_N} 1\{\xi(x+y) = i\}$, $\underline{p} = \min\{p(y) : p(y) > 0\} > 0$, and introduce

$$w_N = 1 - \frac{\|r\|}{\underline{p}} \varepsilon_N \to 1 \text{ as } N \to \infty. \tag{7.5}$$

We use the construction given in (6.2) with other rate functions to provide a coupling of ξ_t^N with some comparison processes. First, let $\xi_t^{N,\mathsf{vm}} \in \{0,1\}^{S_N}$ be the rescaled voter model defined as in (6.2) but using the rate function

$$N w_N c^{N,\mathsf{vm}}(x,\xi),\ x \in S_N, \xi \in \{0,1\}^{S_N}.$$

Next, define the 1-biased rate function

$$\bar{c}^{N,b}(x,\xi) = N w_N c^{N,\mathsf{vm}}(x,\xi) + \widehat{\xi}(x)\Big(2 + \underline{p}^{-1}\Big)\|r\|(\log N)^3 n_1^{(N)}(x,\xi). \tag{7.6}$$

Let $\bar{\xi}_t^N$ be the 1-biased voter model constructed with $\bar{c}^{N,b}$ as in (6.2).

We next verify (6.3) for c^N and $\bar{c}^{N,b}$, that is we will show that for $\xi \le \bar{\xi}$,

$$\begin{aligned} \bar{c}^{N,b}(x,\bar{\xi}) &\ge c^N(x,\xi) \text{ if } \xi(x) = \bar{\xi}(x) = 0, \\ \bar{c}^{N,b}(x,\bar{\xi}) &\le c^N(x,\xi) \text{ if } \xi(x) = \bar{\xi}(x) = 1. \end{aligned} \tag{7.7}$$

First combine (3.23), (3.27) and (3.29) to that for $x \in S_N$ and $\xi \in \{0,1\}^{S_N}$,

$$\begin{aligned}
c^N(x,\xi) &= Nc^{N,\mathsf{vm}}(x,\xi) + \widehat{\xi}(x)\Big[\sum_{\emptyset\neq A\subset\mathcal{N}_N}(\log N r^{N,a}(\sqrt{N}A) + (\log N)^3 r^{N,s}(\sqrt{N}A))1(\xi|_{x+\mathcal{N}_N} = 1_{x+A})\Big]\\
&\quad + \xi(x)\Big[\sum_{\emptyset\neq A\subset\mathcal{N}_N}(\log N)^3 r^{N,s}(\sqrt{N}A)1(\xi|_{x+\mathcal{N}_N} = 1_{x+\mathcal{N}_N\setminus A})\Big]\\
&= Nw_N c^{N,\mathsf{vm}}(x,\xi) + \widehat{\xi}(x)\Big[\sum_{\emptyset\neq A\subset\mathcal{N}_N}(\log N r^{N,a}(\sqrt{N}A) + (\log N)^3 r^{N,s}(\sqrt{N}A))1(\xi|_{x+\mathcal{N}_N} = 1_{x+A})\\
&\qquad + N(1-w_N)f_1^{(N)}(x,\xi)\Big]\\
&\quad + \xi(x)\Big[\sum_{\emptyset\neq A\subset\mathcal{N}_N}(\log N)^3 r^{N,s}(\sqrt{N}A)1(\xi|_{x+\mathcal{N}_N} = 1_{x+\mathcal{N}_N\setminus A}) + N(1-w_N)f_0^{(N)}(x,\xi)\Big]\\
&:= Nw_N c^{N,\mathsf{vm}}(x,\xi) + \widehat{\xi}(x)\tilde{c}_1^N(x,\xi) + \xi(x)\tilde{c}_0^N(x,\xi).
\end{aligned}\tag{7.8}$$

If $\xi|_{x+\mathcal{N}_N} = 1_{x+\mathcal{N}_N\setminus A}$ and $p(\sqrt{N}A) > 0$, then

$$\frac{f_0^{(N)}(x,\xi)}{\underline{p}} = \sum_{y\in\sqrt{N}A}\frac{p(y)}{\underline{p}} \geq 1.$$

Use Lemma 3.7 and then the above to see that

$$\begin{aligned}
\tilde{c}_0^N(x,\xi) &\geq \sum_{\emptyset\neq A\subset\mathcal{N}_N, p(\sqrt{N}A)>0}(\log N)^3(r^{N,s}(\sqrt{N}A)\wedge 0)1(\xi|_{x+\mathcal{N}_N} = 1_{x+\mathcal{N}_N\setminus A}) + \frac{\|r\|}{\underline{p}}(\log N)^3 f_0^{(N)}(x,\xi)\\
&\geq -(\log N)^3\|r\|\frac{f_0^{(N)}(x,\xi)}{\underline{p}} + \frac{\|r\|}{\underline{p}}(\log N)^3 f_0^{(N)}(x,\xi) = 0.
\end{aligned}\tag{7.9}$$

More simply we have

$$\tilde{c}_1^N(x,\xi) \leq \Big[2\|r\|(\log N)^3 + \frac{\|r\|}{\underline{p}}(\log N)^3\Big]n_1^{(N)}(x,\xi).\tag{7.10}$$

Turning to (7.7), assume now that $\xi \leq \bar{\xi}$. If $\xi(x) = \bar{\xi}(x) = 1$, then by (7.8),(7.9) and the monotonicity of the voter model,

$$\begin{aligned}
c^N(x,\xi) &= Nw_N c^{N,\mathsf{vm}}(x,\xi) + \tilde{c}_0^N(x,\xi)\\
&\geq Nw_N c^{N,\mathsf{vm}}(x,\xi) \geq Nw_N c^{N,\mathsf{vm}}(x,\bar{\xi}) = \bar{c}^{N,b}(x,\bar{\xi}),
\end{aligned}$$

the last by (7.6). If $\xi(x) = \bar{\xi}(x) = 0$, then by (7.8) and (7.10),

$$\begin{aligned}
c^N(x,\xi) &= Nw_N f_1^{(N)}(x,\xi) + \tilde{c}_1^N(x,\xi)\\
&\leq Nw_N f_1^{(N)}(x,\bar{\xi}) + (2+\underline{p}^{-1})|r\|(\log N)^3 n_1^{(N)}(x,\xi) = \bar{c}^{N,b}(x,\bar{\xi}),
\end{aligned}$$

where (7.6) is again used in the last equality. This proves (7.7).

More simply (7.7) also holds if $c^N(x,\xi)$ is replaced with $w_N N c^{N,\mathsf{vm}}(x,\xi)$. This is immediate from the monotonicity of the voter model and (7.6). Having verified (6.3) for two pairs of processs we may apply the coupling result (6.4) and conclude that if the three processes $\xi^N, \xi^{N,\mathsf{vm}}, \bar{\xi}^N$ have the same initial state ξ_0^N, then with probability one,

$$\xi_t^N \leq \bar{\xi}_t^N, \text{ and } \xi_t^{N,\mathsf{vm}} \leq \bar{\xi}_t^N \text{ for all } t \geq 0.\tag{7.11}$$

Use these processes to define the empirical processes of one's, $\bar{X}_t^N$ and $X_t^{N,\mathtt{vm}}$ respectively, as in (6.5). We will need to compare these processes over small time periods. We assume u_N satisfies

$$\frac{C_{7.12}}{\sqrt{N}} \le u_N \le (\log N)^{-p} \text{ for some } C_{7.12} > 0, p > 6, \tag{7.12}$$

and recalling that as $N \ge e^3$, we have $u_N(\log N)^3 < 1$.

Lemma 7.1. *For some universal $C_{7.13}$ and all N, if $\xi_0^N = \xi_0^{N,\mathtt{vm}} = \xi_0^N$ then*

$$\begin{aligned} E\big[\bar{X}_{u_N}^N(\mathbf{1}) - X_{u_N}^N(\mathbf{1})\big] &\le C_{7.13}(\log N)^{3-p} X_0^N(\mathbf{1}), \\ E\big[\bar{X}_{u_N}^N(\mathbf{1}) - X_{u_N}^{N,\mathtt{vm}}(\mathbf{1})\big] &\le C_{7.13}(\log N)^{3-p} X_0^N(\mathbf{1}), \\ E\big[\big|X_{u_N}^{N,\mathtt{vm}}(\mathbf{1}) - X_{u_N}^N(\mathbf{1})\big|\big] &\le C_{7.13}(\log N)^{3-p} X_0^N(\mathbf{1}). \end{aligned} \tag{7.13}$$

Proof. For the first inequality, by Lemma 4.1 in [10], $E\bar{X}_s^N(\mathbf{1}) \le e^{(2+\underline{p}^{-1})\|r\|(\log N)^3 s} X_0^N(\mathbf{1})$. Set $\Phi = \mathbf{1}$ in (6.10) to get

$$E[X_t^N(\mathbf{1})] = X_0^N(\mathbf{1}) + \int_0^t E[d^{N,2}(s,\xi_s^N,\mathbf{1}) + d^{N,3}(s,\xi_s^N,\mathbf{1})]\,ds,$$

where by (7.4), $E\big[|d^{N,2}(s,\xi_s^N,\mathbf{1})| + |d^{N,3}(s,\xi_s^N,\mathbf{1})|\big] \le 2\|r\||\mathcal{N}|(\log N)^3 E[X_s^N(\mathbf{1})]$. An elementary integration by parts now implies $E[X_s^N(\mathbf{1})] \ge e^{-2\|r\||\mathcal{N}|(\log N)^3 s} X_0^N(\mathbf{1})$. The above inequalities with $s = u_N$ give the first inequality. The second is even simpler since then $X^{N,\mathtt{vm}}(\mathbf{1})$ is a martingale. The final inequality then follows by the triangle inequality. □

Let $\Phi \in C_b([0,T]\times\mathbb{R}^2)$ and $\|\Phi\|_\infty$ denote its sup norm. Define $|\Phi|_{\mathtt{Lip}}$, respectively $|\Phi|_{1/2,N}$, to be the smallest element in $[0,\infty]$ such that

$$\begin{cases} |\Phi(s,x) - \Phi(s,y)| \le |\Phi|_{\mathtt{Lip}}|x-y|, \quad \forall s \in [0,T], x,y \in \mathbb{R}^2, \\ |\Phi(s-u_N,x) - \Phi(s,x)| \le |\Phi|_{1/2,N}\sqrt{u_N}, \quad \forall s \in [u_N,T], x \in \mathbb{R}^2. \end{cases} \tag{7.14}$$

We will write $\|\Phi\|_{\mathtt{Lip}} = \|\Phi\|_\infty + |\Phi|_{Lip}$, $\|\Phi\|_{1/2,N} = \|\Phi\|_\infty + |\Phi|_{1/2,N}$, and $\|\Phi\|_N = \|\Phi\|_\infty + |\Phi|_{Lip} + |\Phi|_{1/2,N}$. We also will abuse notation slightly and write $\|\Phi_s\|_{\mathtt{Lip}} = \|\Phi_s\|_\infty + |\Phi_s|_{\mathtt{Lip}}$ for the usual Lipschitz norm of $x \to \Phi_s(x) = \Phi(s,x)$. Note that since $u_N < \sqrt{u_N}$ for $N \ge e^3$, we have

$$\text{if, in addition, } \dot{\Phi} \in C_b([0,T]\times\mathbb{R}^2), \text{ then } |\Phi|_{1/2,N} \le \|\dot{\Phi}\|_\infty \text{ for all } N \ge e^3. \tag{7.15}$$

We will often suppress the dependence on N in the above "norms".

We claim that for $\xi \le \eta$, $j = 2,3$,

$$|d^{N,j}(x,\eta) - d^{N,j}(x,\xi)| \le 2\|r\| \sum_{y \in x+\bar{\mathcal{N}}_N} (\eta(y) - \xi(y)). \tag{7.16}$$

If $\xi \ne \eta$ on $x + \bar{\mathcal{N}}_N$, then the right-hand side is at least $2\|r\|$ and so the above follows from (7.3) and the triangle inequality. If $\xi|_{\bar{\mathcal{N}}_N} = \eta|_{\bar{\mathcal{N}}_N}$ then the left-hand side is zero, and so (7.16) is trivial.

Lemma 7.2. *There is a constant $C_{7.17} > 0$ such that for $j = 2,3$, all $T > 0$, $\Phi \in C_b([0,T]\times\mathbb{R}^2)$ and all $s \in [0,T]$,*

$$E_{\xi_0^N}\left[\left|d^{N,j}(s,\xi_{u_N}^N,\Phi) - d^{N,j}(s,\xi_{u_N}^{N,\mathtt{vm}},\Phi)\right|\right] \le C_{7.17}\|\Phi\|_\infty(\log N)^{6-p} X_0^N(\mathbf{1}). \tag{7.17}$$

Proof. Using (7.16) with $\eta = \bar\xi^N_{u_N}$ and $\xi = \xi^N_{u_N}$ and in (6.8) we see that

$$\left| d^{N,j}(s, \bar\xi^N_{u_N}, \Phi) - d^{N,j}(s, \xi^N_{u_N}, \Phi) \right| \le 2\|r\|\|\Phi\|_\infty \frac{\ell_N^{(j)}}{N'} \sum_{x \in S_N} \Big[\sum_{y \in x + \bar{\mathcal{N}}_N} \bar\xi^N_{u_N}(y) - \xi^N_{u_N}(y) \Big]$$
$$= 2\|r\||\bar{\mathcal{N}}|\|\Phi\|_\infty \ell_N^{(j)} \big[\bar X^N_{u_N}(\mathbf{1}) - X^N_{u_N}(\mathbf{1})\big].$$

Similarly,

$$\left| d^{N,j}(s, \bar\xi^N_{u_N}, \Phi) - d^{N,j}(s, \xi^{N,\mathsf{vm}}_{u_N}, \Phi) \right| \le 2\|r\||\bar{\mathcal{N}}|\|\Phi\|_\infty \ell_N^{(j)} \big[\bar X^N_{u_N}(\mathbf{1}) - X^N_{u_N}(\mathbf{1})\big].$$

It follows that the left-hand side of (7.17) is bounded by

$$2\|r\||\bar{\mathcal{N}}|\|\Phi\|_\infty (\log N)^3 E\big[(\bar X^N_{u_N}(\mathbf{1}) - X^N_{u_N}(\mathbf{1})) + (\bar X^N_{u_N}(\mathbf{1}) - X^{N,\mathsf{vm}}_{u_N}(\mathbf{1}))\big].$$

Now Lemma 7.1 completes the proof. □

The next step is to consider $E[d^{N,j}(s, \xi^{N,\mathsf{vm}}_{u_N}, \Phi)]$. Under a probability $\widehat P$, let $\{B^{N,x}_t, x \in S_N\}$ be a rate $w_N N$, coalescing random walk system on S_N with jump kernel p_N and for $A \subset S_N$ let $B^{N,A}_t = \cup_{x \in A}\{B^{N,x}_t\}$. For finite nonempty disjoint $A_i \subset S_N$ define $\sigma^N(A_1, \dots, A_n)$ and $\tau^N(A_1, \dots, A_n)$ in the same way as σ and τ are defined in (1.6), but with $\{B^{N,x} : x \in S_N\}$ in place of $\{B^x : x \in \mathbb{Z}^2\}$. Similarly define $\sigma^N(x_1, \dots x_n)$ and $\tau^N(x_1, \dots, x_n)$ for distinct x_i in S_N. If $x \in S_N$, then introduce $\sigma^N_x(A_1, \dots, A_n) = \sigma^N(x + A_1, \dots, x + A_n)$, and similarly for $\tau^N_x(A_1, \dots, A_n)$.

The duality equation connecting the voter model $\xi^{N,\mathsf{vm}}$ with the coalescing random walks $\{B^{N,x}_t\}$ that we need is the following (see Section III.4 of Lig85). For $x \in S_N$, finite disjoint sets $A, B \subset S_N$ and $\xi^N_0 \in \{0,1\}^{S_N}$,

$$E_{\xi^N_0}\Big[\prod_{a \in A} \xi^{N,\mathsf{vm}}_t(a) \prod_{b \in B}(1 - \xi^{N,\mathsf{vm}}_t(b))\Big] = \widehat E\Big[\prod_{a \in A} \xi^N_0(B^{N,a}_t) \prod_{b \in B}(1 - \xi^N_0(B^{N,b}_t))\Big]. \tag{7.18}$$

For $A \subset \mathcal{N}_N$ define

$$\begin{aligned} I^{N,+}(x, u_N, A, \xi^N_0) &= \prod_{a \in A} \xi^N_0(B^{N,x+a}_{u_N}) \prod_{b \in \bar{\mathcal{N}}_N \setminus A}(1 - \xi^N_0(B^{N,x+b}_{u_N})) \\ I^{N,-}(x, u_N, A, \xi^N_0) &= \prod_{a \in \bar{\mathcal{N}}_N \setminus A} \xi^N_0(B^{N,x+a}_{u_N}) \prod_{b \in A}(1 - \xi^N_0(B^{N,x+b}_{u_N})). \end{aligned} \tag{7.19}$$

With this notation, (7.1) and (7.18) imply that

$$\begin{aligned} E_{\xi^N_0}\big(d^{N,2}(x, \xi^{N,\mathsf{vm}}_{u_N})\big) &= \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,a}\sqrt{N}A)\widehat E\big(I^{N,+}(x, u_N, A, \xi^N_0)\big), \\ E_{\xi^N_0}\big(d^{N,3}(x, \xi^{N,\mathsf{vm}}_{u_N})\big) &= \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A)\Big[\widehat E\big(I^{N,+}(x, u_N, A, \xi^N_0)\big) - \widehat E\big(I^{N,-}(x, u_N, A, \xi^N_0)\big)\Big]. \end{aligned} \tag{7.20}$$

If we now define

$$\hat H^{N,2}(\xi^N_0, u_N, \Phi_s) = \frac{\ell_N^{(2)}}{N'} \sum_{x \in S_N} \Phi(s,x) \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,a}(\sqrt{N}A)\widehat E[I^{N,+}(x, u_N, A, \xi^N_0)] \tag{7.21}$$

$$\hat H^{N,3}(\xi^N_0, u_N, \Phi_s) = \frac{\ell_N^{(3)}}{N'} \sum_{x \in S_N} \Phi(s,x) \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A)\Big(\widehat E[I^{N,+}(x, u_N, A, \xi^N_0)] - \widehat E[I^{N,-}(x, u_N, A, \xi^N_0)]\Big), \tag{7.22}$$

then for $j = 2, 3$,

$$E_{\xi^N_0}[d^{N,j}(s, \xi^{N,\mathsf{vm}}_{u_N}, \Phi)] = \hat H^{N,j}(\xi^N_0, u_N, \Phi_s). \tag{7.23}$$

Lemma 7.3. *There is a constant $C_{7.24}$ such for $j=2,3$, all $T>0$, $\Phi\in C_b([0,T]\times\mathbb{R}^2)$ and all $s\in[u_N,T]$,*

$$\Big|E_{\xi_0^N}\big(d^{N,j}(s,\xi_s^N,\Phi)|\mathcal{F}^N_{s-u_N}\big)-\hat H^{N,j}(\xi^N_{s-u_N},u_N,\Phi_{s-u_N})\Big|\le C_{7.24}\|\Phi\|_{1/2,N}(\log N)^{3-\frac{p}{2}}X^N_{s-u_N}(\mathbf{1}). \tag{7.24}$$

Proof. By the Markov property, Lemma 7.2 and (7.23), the left-hand side of (7.24) is bounded above by

$$C_{7.17}\|\Phi\|_\infty(\log N)^{6-p}X^N_{s-u_N}(\mathbf{1})+\Big|\hat H^{N,j}(\xi^N_{s-u_N},u_N,\Phi_s)-\hat H^{N,j}(\xi^N_{s-u_N},u_N,\Phi_{s-u_N})]\Big|. \tag{7.25}$$

Use (7.23) and then (7.2) with the voter duality (7.18) to see that the second term above is bounded by

$$\begin{aligned}&\|r\|\frac{\ell_N^{(j)}}{N'}\sum_{x\in S_N}|\Phi(s,x)-\Phi(s-u_N,x)|\widehat E\Big[\sum_{y\in\bar{\mathcal{N}}_N}\xi^N_{s-u_N}(B^{N,x+y}_{u_N})\Big]\\ &\qquad\le\|r\||\Phi|_{1/2}\sqrt{u}_N\ell_N^{(j)}\sum_{y\in\bar{\mathcal{N}}_N}\widehat E\Big[\frac{1}{N'}\sum_{x\in S_N}\xi^N_{s-u_N}(x+B^{N,y}_{u_N})\Big]\\ &\qquad=\|r\||\Phi|_{1/2}|\bar{\mathcal{N}}|(\log N)^{3-\frac{p}{2}}X^N_{s-u_N}(\mathbf{1}).\end{aligned}$$

Inserting this bound into (7.25) completes the proof. ☐

We next further decompose the $\hat H^{N,j}$. Consider $\emptyset\neq A\subset\mathcal{N}_N$. If $I^{N,\pm}(x,u_N,A,\xi_0^N)\neq 0$ then necessarily $2\le|B^{N,x+\bar{\mathcal{N}}_N}_{u_N}|\le|\bar{\mathcal{N}}|$, and by defining

$$I_i^{N,\pm}(x,u_N,A,\xi_0^N)=I^{N,\pm}(x,u_N,A,\xi_0^N)1\{|B^{N,x+\bar{\mathcal{N}}_N}_{u_N}|=i\}, \tag{7.26}$$

we can write

$$I^{N,\pm}(x,u_N,A,\xi_0^N)=\sum_{i=2}^{|\bar{\mathcal{N}}|}I_i^{N,\pm}(x,u_N,A,\xi_0^N). \tag{7.27}$$

Letting

$$h_{i,j}^{N,\pm}(\xi_0^N,u_N,A,\Phi_s)=\frac{\ell_N^{(j)}}{N'}\sum_{x\in S_N}\Phi(s,x)I_i^{N,\pm}(x,u_N,A,\xi_0^N) \tag{7.28}$$

and

$$\begin{aligned}\hat H_i^{N,2}(\xi_0^N,u_N,\Phi_s)&=\sum_{\emptyset\neq A\subset\mathcal{N}_N}r^{N,a}(\sqrt{N}A)\widehat E\Big(h_{i,2}^{N,+}(\xi_0^N,u_N,A,\Phi_s)\Big)\\ \hat H_i^{N,3}(\xi_0^N,u_N,\Phi_s)&=\sum_{\emptyset\neq A\subset\mathcal{N}_N}r^{N,s}(\sqrt{N}A)\widehat E\Big(h_{i,3}^{N,+}(\xi_0^N,u_N,A,\Phi_s)-h_{i,3}^{N,-}(\xi_0^N,u_N,A,\Phi_s)\Big),\end{aligned} \tag{7.29}$$

we obtain the decomposition

$$\hat H^{N,j}(\xi_0^N,u_N,\Phi_s)=\sum_{i=2}^{|\bar{\mathcal{N}}|}\hat H_i^{N,j}(\xi_0^N,u_N,\Phi_s). \tag{7.30}$$

We will now obtain simple bounds on the summands in (7.30), leaving a more detailed analysis of the main terms $\hat H_2^{N,2}(\xi_0^N,u_N,\Phi_s)$ and $\hat H_3^{N,3}(\xi_0^N,u_N,\Phi_s)$ to Section 7.3.

Lemma 7.4. *There are constants $C_{7.31}$, $C_{7.32}$ such that for any $T>0$, $\phi \in C_b([0,T]\times\mathbb{R}^2)$ and $s\in[0,T]$,*

$$|\hat H_i^{N,2}(\xi_0^N,u_N,\Phi_s)| \le C_{7.31}\|r\|\|\Phi_s\|_\infty(\log N)^{1-\binom{i}{2}}X_0^N(\mathbf{1}), \quad 2\le i\le|\bar{\mathcal{N}}|, \tag{7.31}$$

$$|\hat H_i^{N,3}(\xi_0^N,u_N,\Phi_s)| \le C_{7.32}\|r\|\|\Phi_s\|_\infty(\log N)^{3-\binom{i}{2}}X_0^N(\mathbf{1}), \quad 3\le i\le|\bar{\mathcal{N}}|, \tag{7.32}$$

and

$$|\hat H_2^{N,3}(\xi_0^N,u_N,\Phi_s)| \le C_{7.32}\|r\|\|\Phi_s\|_{Lip}(\log N)^{(6-p)/2}X_0^N(\mathbf{1}). \tag{7.33}$$

Proof. We start with the observation that for $\emptyset\ne A\subset\mathcal{N}_N$ and any $a\in A$,

$$|I_i^{N,\pm}(x,u_N,A,\xi_0^N)| \le \xi_0^N(B_{u_N}^{N,x+a})1\{|B_{u_N}^{N,x+\bar{\mathcal{N}}_N}|=i\} + \xi_0^N(B_{u_N}^{N,x})1\{|B_{u_N}^{N,x+\bar{\mathcal{N}}_N}|=i\}. \tag{7.34}$$

We will bound $E\big(|h_{i,j}^{N,\pm}|\big)$ with this inequality. By translation invariance,

$$\begin{aligned}
&\frac{\ell_N^{(j)}}{N'}\sum_{x\in S_N}|\Phi(s,x)|\widehat E\big(\xi_0^N(B_{u_N}^{N,x+a})1\{|B_{u_N}^{N,x+\bar{\mathcal{N}}_N}|=i\}\big)\\
&\quad=\frac{\ell_N^{(j)}}{N'}\sum_{x\in S_N}\sum_{w\in S_N}|\Phi(s,x)|\xi_0^N(x+w)\widehat P\big(B_{u_N}^{N,x+a}=x+w,|B_{u_N}^{N,x+\bar{\mathcal{N}}_N}|=i\big)\\
&\quad=\frac{\ell_N^{(j)}}{N'}\sum_{x\in S_N}\sum_{w\in S_N}|\Phi(s,x)|\xi_0^N(x+w)\widehat P\big(B_{u_N}^{N,a}=w,|B_{u_N}^{N,\bar{\mathcal{N}}_N}|=i\big)\\
&\quad\le\|\Phi_s\|_\infty X_0^N(\mathbf{1})\ell_N^{(j)}\widehat P\big(|B_{u_N}^{N,\bar{\mathcal{N}}_N}|=i\big).
\end{aligned}$$

The same bound holds if we replace $B_{u_N}^{N,x+a}$ with $B_{u_N}^{N,x}$, and thus we have shown that

$$\widehat E\big(|h_{ij}^{N,\pm}(\xi_0^N,u_N,A,\Phi_s)|\big)\le 2\|\Phi_s\|_\infty X_0^N(\mathbf{1})\ell_N^{(j)}\widehat P\big(|B_{u_N}^{N,\bar{\mathcal{N}}_N}|=i\big). \tag{7.35}$$

By Remark 1.7, the lower bound on u_N in (7.12), and the fact that $w_N\ge 1/2$ for N large,

$$\ell_N^{(j)}\widehat P\big(|B_{u_N}^{N,\bar{\mathcal{N}}_N}|=i\big)=\begin{cases}O((\log N)^{1-\binom{i}{2}}), & j=2,\\ O((\log N)^{3-\binom{i}{2}}), & j=3.\end{cases}$$

Using this bound in (7.35) above, and substituting into (7.29) we obtain (7.31) and (7.32).

The proof of (7.33) is more delicate, as it relies on cancellation. For $x\in S_N$ define

$$\Omega_x^N(A)=\{\sigma_x^N(A,\bar{\mathcal{N}}_N\setminus A)>u_N,\tau_x^N(A,\bar{\mathcal{N}}_N\setminus A)<u_N\}. \tag{7.36}$$

If $|B_{u_N}^{N,\bar{\mathcal{N}}_N}|=2$, and $B_{u_N}^{N,x+A}$ and $B_{u_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}$ are disjoint, then $B_{u_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}=B_{u_N}^{N,x}$ and for any $a\in A$, $B_{u_N}^{N,x+A}=B_{u_N}^{N,x+a}$. Thus

$$\begin{aligned}
&I_2^{N,+}(x,u_N,A,\xi_0^N)-I_2^{N,-}(x,u_N,A,\xi_0^N)\\
&\quad=\Big(\xi_0^N(B_{u_N}^{N,x+a})(1-\xi_0^N(B_{u_N}^{N,x}))-(1-\xi_0^N(B_{u_N}^{N,x+a}))\xi_0^N(B_{u_N}^{N,x})\Big)1\{\Omega_x^N(A)\}\\
&\quad=\Big(\xi_0^N(B_{u_N}^{N,x+a})-\xi_0^N(B_{u_N}^{N,x}\Big)1\{\Omega_x^N(A)\}\\
&\quad=\sum_{w\in S_N}\xi_0^N(w)1\{B_{u_N}^{N,x+a}=w\}1\{\Omega_x^N(A)\}-\sum_{w\in S_N}\xi_0^N(w)1\{B_{u_N}^{N,x}=w\}1\{\Omega_x^N(A)\}.
\end{aligned}$$

Now by translation invariance,

$$\begin{aligned}
&\widehat E[I_2^{N,+}(x,u_N,A,\xi_0^N)-I_2^{N,-}(x,u_N,A,\xi_0^N)]\\
&=\sum_{w\in S_N}\xi_0^N(w)\widehat E[1\{B_{u_N}^{N,a}=w-x\}1\{\Omega_0^N(A)\}]-\sum_{w\in S_N}\xi_0^N(w)\widehat E[1\{B_{u_N}^{N,0}=w-x\}1\{\Omega_0^N(A)\}].
\end{aligned}$$

Plugging into the definition of $h_{2,3}^{N,\pm}(\xi_0^N, u_N, A, \Phi_s)$ gives

$$\begin{aligned}\widehat{E}\big((h_{2,3}^{N,+} - h_{2,3}^{N,-})(\xi_0^N, u_N, A, \Phi_s)\big) &= \frac{(\log N)^3}{N'} \sum_{w\in S_N} \xi_0^N(w) \sum_{x\in S_N} \Phi(s,x)\Big[\widehat{E}\big(1\{B_{u_N}^{N,a} = w - x\}1\{\Omega_0^N(A)\}\big) \\ &\qquad - \widehat{E}\big(1\{B_{u_N}^{N,0} = w - x\}1\{\Omega_0^N(A)\}\big)\Big] \\ &= \frac{(\log N)^3}{N'} \sum_{w\in S_N} \xi_0^N(w)\widehat{E}\Big[(\Phi(s, w - B_{u_N}^{N,a}) - \Phi(s, w - B_{u_N}^{N,0}))1\{\Omega_0^N(A)\}\Big].\end{aligned}$$

By the above,

$$\begin{aligned}\Big|\widehat{E}\big((h_{2,3}^{N,+} - h_{2,3}^{N,-})(\xi_0^N, u_N, A, \Phi_s))\big]\Big| &\le \frac{(\log N)^3}{N'} \sum_{w\in S_N} \xi_0^N(w)\widehat{E}\big(\big|\Phi(s, w - B_{u_N}^{N,a}) - \Phi(s, w - B_{u_N}^{N,0})\big|\big) \\ &\le \|\Phi_s\|_{\mathtt{Lip}} \frac{(\log N)^3}{N'} \sum_{w} \xi_0^N(w)\widehat{E}\big(|B_{u_N}^{N,a} - B_{u_N}^{N,0}|\big) \\ &= \|\Phi_s\|_{\mathtt{Lip}} X_0^N(\mathbf{1})(\log N)^3 \widehat{E}\big(|a| + |B_{2u_N}^{N,0}|\big) \\ &\le \|\Phi_s\|_{\mathtt{Lip}} X_0^N(\mathbf{1})(\log N)^3 \big(c/\sqrt{N} + \sqrt{\sigma^2 4u_N}\big) \\ &\le C(\mathcal{N}, \sigma^2)\|\Phi_s\|_{\mathtt{Lip}} X_0^N(\mathbf{1})(\log N)^{(6-p)/2}, \end{aligned} \tag{7.37}$$

where $c = \max_{e\in\mathcal{N}} |e|$ and we recall from (3.13) that $\widehat{E}\big(|B_s^{N,0}|^2\big) = 2w_N\sigma^2 s \le 2\sigma^2 s$. Using this bound in (7.29) we obtain (7.33). □

In order to use the above results to effectively handle the drift terms $d^{N,2}(x, \xi_s^N, \Phi)$ and $d^{N,3}(x, \xi_s^N, \Phi)$ we must first obtain bounds on the first and second moments of the total mass, which will play an important roll in what follows. Therefore we interrupt our current analysis to handle the total mass next.

7.2 Total mass bounds

We now introduce a particular choice of u_N, namely

$$t_N = (\log N)^{-19}. \tag{7.38}$$

Lemma 7.5. *There is a constant $C_{7.40} > 0$ so that for $j = 2, 3$, all $T > 0$ and $\Phi \in C_b([0, T] \times \mathbb{R}^2)$,*

$$|d^{N,j}(x, \xi_s^N, \Phi)| \le \|r\|\|\Phi\|_\infty |\bar{\mathcal{N}}| \ell_N^{(j)} X_s^N(\mathbf{1}) \quad \forall s \in [0, T] \tag{7.39}$$

and

$$\Big|E[d^{N,j}(s, \xi_s^N, \Phi)|\mathcal{F}_{s-t_N}^N]\Big| \le C_{7.40}\|\Phi\| X_{s-t_N}^N(\mathbf{1}) \quad \forall s \in [t_N, T]. \tag{7.40}$$

Proof. (7.39) holds by (7.4), while (7.40) follows from Lemma 7.3, (7.30), and Lemma 7.4. □

Proposition 7.6. *There exists a $c_{7.41} > 0$, and for $T > 0$ a constant $C_{7.41} > 0$ depending on T, such that for any $t \le T$,*

$$\begin{aligned} &(a) \qquad E[X_t^N(\mathbf{1})] \le (1 + C_{7.41}(\log N)^{-16}) X_0^N(\mathbf{1}) \exp(c_{7.41} t), \\ &(b) \qquad E[(X_t^N(\mathbf{1}))^2] \le C_{7.41}(X_0^N(\mathbf{1}) + (X_0^N(\mathbf{1}))^2). \end{aligned} \tag{7.41}$$

Therefore, for $T > 0$ there is a constant $C_{7.42} > 0$, depending on T, such that for all $s, t \in [0, T]$,

$$E[X_s^N(\mathbf{1})X_t^N(\mathbf{1})] \le C_{7.42}(X_0^N(\mathbf{1}) + (X_0^N(\mathbf{1}))^2). \tag{7.42}$$

Proof of (a). By (6.10),

$$E[X_t^N(\mathbf{1})] = X_0^N(\mathbf{1}) + \int_0^{t\wedge t_N} \sum_{j=2}^3 E[d^{N,j}(s, \xi_s^N, \mathbf{1})]ds + \int_{t\wedge t_N}^t \sum_{j=2}^3 E\Big[E[d^{N,j}(s, \xi_s^N, \mathbf{1})|\mathcal{F}_{s-t_N}^N]\Big]ds. \tag{7.43}$$

Use the two bounds from Lemma 7.5 (with $\Phi = 1$) and $\ell_N^{(j)} \le (\log N)^3$ in (7.43) to get

$$E[X_t^N(\mathbf{1})] \le X_0^N(\mathbf{1}) + 2\|r\||\bar{\mathcal{N}}|(\log N)^3 \int_0^{t\wedge t_N} E[X_s^N(\mathbf{1})]ds + C_{7.40} \int_0^{(t-t_N)^+} E[X_s^N(\mathbf{1})]ds. \tag{7.44}$$

This implies that for $t \le t_N$,

$$E[X_t^N(\mathbf{1})] \le X_0^N(\mathbf{1}) + 2\|r\||\bar{\mathcal{N}}|(\log N)^3 \int_0^t E[X_s^N(\mathbf{1})]ds, \tag{7.45}$$

and thus by Gronwall's inequality, for $t \le t_N$,

$$E[X_t^N(\mathbf{1})] \le \exp(2\|r\||\bar{\mathcal{N}}|(\log N)^3 t)X_0^N(\mathbf{1}) \le \exp(2\|r\||\bar{\mathcal{N}}|(\log N)^{-16})X_0^N(\mathbf{1}) \le e^{2\|r\||\bar{\mathcal{N}}|}X_0^N(\mathbf{1}). \tag{7.46}$$

By plugging this bound into (7.44) we obtain for all $t > 0$, (recall $N \ge 3$)

$$E[X_t^N(\mathbf{1})] \le X_0^N(\mathbf{1}) + e^{2\|r\||\bar{\mathcal{N}}|}2\|r\||\bar{\mathcal{N}}|(\log N)^{-16}X_0^N(\mathbf{1}) + C_{7.40}\int_0^t E[X_s^N(\mathbf{1})]ds.$$

Another use of Gronwall's inequality completes the proof of part (a). □

Before proving (b) we establish some preparatory results which will also be useful later. Let

$$K_N = \sum_{y\in\mathbb{N}_N} p_N(y) \log N \widehat{P}(\sigma^N(0, y) > t_N). \tag{7.47}$$

By Lemma A.3(ii) in [6],

$$\lim_{N\to\infty} K_N = 2\pi\sigma^2. \tag{7.48}$$

From (6.10) we have

$$E[X_t^N(\mathbf{1})^2] \le 4\Big[X_0^N(\mathbf{1})^2 + E[\langle M^N(\mathbf{1})\rangle_t] + E[D_t^{N,2}(\mathbf{1})^2] + E[D_t^{N,3}(\mathbf{1})^2]\Big]. \tag{7.49}$$

Recall from (6.11) that $\langle M^N(\Phi)\rangle_t = \langle M^N(\Phi)\rangle_{1,t} + \langle M^N(\Phi)\rangle_{2,t}$.

Lemma 7.7. *For any $T > 0$ there is a constant $C_{7.50} > 0$, depending on T, so that for any bounded Borel function Φ on $[0, T] \times \mathbb{R}^2$ and all $t \in [0, T]$,*

$$\langle M^N(\Phi)\rangle_{1,t} = 2\int_0^t X_s^N((\log N)\Phi_s^2 f_0^{(N)}(\cdot, \xi_s^N))ds + \int_0^t \tilde{m}_{1,s}^N(\Phi)ds$$
$$\text{where } |\tilde{m}_{1,s}^N(\Phi)| \le C_{7.50}\frac{\|\Phi\|_{Lip}^2(\log N)}{\sqrt{N}}X_s^N(\mathbf{1}), \tag{7.50}$$

and

$$\langle M^N(\Phi)\rangle_{2,t} = \int_0^t \tilde{m}_{2,s}^N(\Phi)ds, \text{ where } |\tilde{m}_{2,s}^N(\Phi)| \le 2\|r\||\bar{\mathcal{N}}|\|\Phi\|_\infty^2 \frac{\ell_N^{(3)}}{N'} X_s^N(\mathbf{1})ds. \tag{7.51}$$

Proof. (7.50) holds by the corresponding result in Lemma 4.8 of [12]. Note that the result there is actually an identity and bound for a generic state $\xi \in \{0,1\}^{S_N}$ (although it is not stated as such) and applies immediately to our setting as well. For (7.51) note first that by their definitions, $|c^{N,a}(x,\xi)| = |d^{N,2}(x,\xi)|$ and $|c^{N,s}(x,\xi)| = |d^{N,3}(x,\xi)|$. The second result now follows immediately by using (7.4) to bound the absolute value of the integrand of the integrals in (6.12) for $\langle M^N(\Phi)\rangle_{2,t}$. □

Lemma 7.8. *There is a $C_{7.54} > 0$ so that for any bounded Borel Φ on $\mathbb{R}^2$:*
(a)

$$\frac{\log N}{N'} \sum_{x\in S_N} \sum_{y\in\mathcal{N}_N} \Phi(x) p_N(y) \hat{E}(\xi_0^N(B^{N,x}_{t_N}) 1(\sigma^N(x,x+y) > t_N)) = K_N X_0^N(\Phi) + \mathcal{E}_N, \quad (7.52)$$

where

$$|\mathcal{E}_N| \le C_{7.54} |\Phi|_{Lip} (\log N)^{-17/2} X_0^N(\mathbf{1}). \quad (7.53)$$

(b) If, in addition, $\Phi \ge 0$, then

$$E\Big[X^N_{t_N}((\log N)\Phi f_0^{(N)}(\cdot, \xi^N_{t_N}))\Big] \le C_{7.54}\Big(\|\Phi\|_{Lip}(\log N)^{-17/2} X_0^N(\mathbf{1}) + X_0^N(\Phi)\Big). \quad (7.54)$$

Proof. (a) Let Σ_N denote the left-hand side of (7.52). We may write $B^{N,x}$ as $x + B^{N,0}$ and sum over the values of $B^{N,0}_{t_N}$ to see that

$$\begin{aligned}\Sigma_N &= \frac{\log N}{N'} \sum_{x\in S_N}\sum_{w\in S_N}\sum_{y\in\mathcal{N}_N} p_N(y)[\Phi(x) - \Phi(x+w)]\xi_0^N(x+w)\hat{P}(B^{N,0}_{t_N} = w, \sigma^N(0,y) > t_N) \\ &\quad + \frac{\log N}{N'} \sum_{x\in S_N}\sum_{w\in S_N} \Phi(x+w)\xi_0^N(x+w) \sum_{y\in\mathcal{N}_N} p_N(y)\hat{P}(B^{N,0}_{t_N} = w, \sigma^N(0,y) > t_N) \\ &:= \Sigma_N^{(1)} + \Sigma_N^{(2)}.\end{aligned}$$

Use the implicit Lipschitz continuity of Φ to see that

$$\begin{aligned}|\Sigma_N^{(1)}| &\le |\Phi|_{\mathtt{Lip}} \sum_{w\in S_N} |w| \Big(\frac{\log N}{N'}\sum_{x\in S_N} \xi_0^N(x+w)\Big) \sum_{y\in\mathcal{N}_N} p_N(y)\hat{P}(B^{N,0}_{t_N} = w, \sigma^N(0,y) > t_N) \\ &\le |\Phi|_{\mathtt{Lip}} X_0^N(\mathbf{1}) \log N \hat{E}(|B^{N,0}_{t_N}|) \le |\Phi|_{\mathtt{Lip}} X_0^N(\mathbf{1}) \log N \sqrt{2}\sigma\sqrt{t_N}. \quad (7.55)\end{aligned}$$

The fact that $w_N \le 1$ is used in the last line as well. Sum over x first and then w, to see that $\Sigma_N^{(2)} = K_N X_0^N(\Phi)$. This gives (a).

(b) Assume now $\Phi \ge 0$. We will compare with the corresponding expression for $\xi^{N,\mathtt{vm}}_{t_N}$, and use the duality (7.18) to compute the latter, and hence see that the left-hand side of (7.54) equals

$$\begin{aligned}&\frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}} \Phi(x)p_N(y) E_{\xi_0^N}[\xi^N_{t_N}(x)\widehat{\xi}^N_{t_N}(x+y) - \xi^{N,\mathtt{vm}}_{t_N}(x)\widehat{\xi}^{N,\mathtt{vm}}_{t_N}(x+y)] \\ &\qquad + \frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}_N} \Phi(x)p_N(y)\hat{E}[\xi_0^N(B^{N,x}_{t_N})\widehat{\xi}_0^N(B^{N,x+y}_{t_N})]. \quad (7.56)\end{aligned}$$

By the coupling (7.11), Lemma 7.1, and the triangle inequality, the absolute value of the first sum in (7.56) is bounded by

$$\begin{aligned}&\|\Phi\|_\infty \frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}} p_N(y)\Big(E[\bar{\xi}^N_{t_N}(x) - \xi^N_{t_N}(x)] + E[\bar{\xi}^N_{t_N}(x) - \xi^{N,\mathtt{vm}}_{t_N}(x)] \\ &\qquad + E[\bar{\xi}^N_{t_N}(x+y) - \xi^N_{t_N}(x+y)] + E[\bar{\xi}^N_{t_N}(x+y) - \xi^{N,\mathtt{vm}}_{t_N}(x+y)]\Big) \\ &\qquad = 2\|\Phi\|_\infty(\log N)\Big(E[\bar{X}^N_{t_N}(\mathbf{1}) - X^N_{t_N}(\mathbf{1})] + E[\bar{X}^N_{t_N}(\mathbf{1}) - X^{N,\mathtt{vm}}_{t_N}(\mathbf{1})]\Big) \\ &\qquad \le C\|\Phi\|_\infty(\log N)^{-15} X_0^N(\mathbf{1}). \quad (7.57)\end{aligned}$$

The second sum in (7.56) is bounded by Σ_N (the left-hand side of part (a)), and so (a) and (7.48) give the required bound. □

Proof of Proposition 7.6(b). It now follows quite easily from Lemmas 7.7 and 7.8(b), as well as part (a), that there is a constant $C_{7.58}$ depending on T such that,

$$E[\langle M^N(\mathbf{1})\rangle_T] \le C_{7.58} X_0^N(\mathbf{1}). \tag{7.58}$$

Here one uses the Markov property and Lemma 7.8(b) to bound $\int_{t_N}^T X_s^N((\log N) f_0^{(N)}(\xi_s^N))ds$, and the obvious crude bound on the integrand to handle the integral from 0 to t_N.

Turning to the last terms in (7.49), we consider more generally $E[(D_t^{N,j}(\Phi))^2]$, where $\Phi : \mathbb{R}^2 \to \mathbb{R}$ is Lipschitz continuous. We claim that for $j = 2, 3$,

$$E[(D^{N,j}_{t\wedge t_N}(\Phi))^2] \le \|\Phi\|_\infty^2 |\bar{\mathcal{N}}|^2 (\log N)^6 t_N E\Big[\int_0^{t\wedge t_N} X_s^N(\mathbf{1})^2 ds\Big], \tag{7.59}$$

and there is a constant $C_{7.60} > 0$ such that if $t_2 > t_1 \ge t_N$, then

$$\begin{aligned} &E[(D^{N,j}_{t_2}(\Phi) - D^{N,j}_{t_1}(\Phi))^2] \\ &\quad \le C_{7.60}\Big(\|\Phi\|^2_{\texttt{Lip}}(t_2 - t_1) + \|\Phi\|^2_\infty (\log N)^6 (t_N \wedge (t_2 - t_1))\Big) \int_{t_1}^{t_2} E[X_s^N(\mathbf{1})^2] ds. \end{aligned} \tag{7.60}$$

The proofs of the corresponding inequalities (66) and (69) in [9] apply here without change if (41) and (48) there are replaced by (7.39) and (7.40) here, using $\ell_N^{(2)} = \log N \le (\log N)^3$. (The increment bound is stronger than we need here but will be useful later in establishing tightness.) Now use (7.58), (7.59) and (7.60) in (7.49) to see there is a constant $C_{7.61}$ depending on T such that if $t \le T$ then

$$E[X_t^N(\mathbf{1})^2] \le C_{7.61}\Big[X_0^N(\mathbf{1})^2 + X_0^N(\mathbf{1}) + \int_0^t E[X_s^N(\mathbf{1})^2] ds\Big]. \tag{7.61}$$

A simple Gronwall argument finishes the proof of Proposition 7.6 (b). □

7.3 Exact drift asymptotics and first moment measure bounds

Introduce, for $\delta > 0$,

$$\mathscr{I}^N(\delta, \xi_0^N) = \int\int 1_{\{|w-z|<\delta\}} dX_0^N(w) dX_0^N(z), \tag{7.62}$$

and define

$$\mathscr{I}^N(\xi_0^N) = \mathscr{I}^N\Big(\sqrt{t_N(\log N)^{1/2}}, \xi_0^N\Big). \tag{7.63}$$

Lemma 7.9. *There is a $C_{7.64}$ so that for $N \ge N(\varepsilon_0)$, distinct $a_0, a_1, a_2 \in \bar{\mathcal{N}}_N$, and $\xi_0^N \in \{0,1\}^{S_N}$,*

$$\begin{aligned} &\frac{(\log N)^3}{N'} \sum_{x \in S_N} \widehat{P}\Big(\xi_0^N(B_{t_N}^{N,x+a_1}) = \xi_0^N(B_{t_N}^{N,x+a_2}) = 1, \sigma_x^N(a_0,a_1,a_2) > t_N\Big) \\ &\qquad \le C_{7.64}\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-(1/2)} X_0^N(\mathbf{1})\Big), \end{aligned} \tag{7.64}$$

and

$$\begin{aligned} &\frac{(\log N)}{N'} \sum_{x \in S_N} \widehat{P}\Big(\xi_0^N(B_{t_N}^{N,x+a_1}) = \xi_0^N(B_{t_N}^{N,x+a_2}) = 1, \sigma_x^N(a_1,a_2) > t_N\Big) \\ &\qquad \le C_{7.64}\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-(1/2)} X_0^N(\mathbf{1})\Big). \end{aligned} \tag{7.65}$$

Proof. The proof of (7.64) may be found in the derivation of (46) in [9] with $\Phi = \mathbf{1}$. See, in particular, the bound on $\mathcal{T}_N(\Phi)$ in that proof on pages 1214 and 1215, and note we are setting $\eta = 1/2$ in the final display of Section 4.2 of that article. In that result the stochastic process $\xi_u^N(\mathbf{1})$ is a rescaled two-dimensional Lotka-Volterra model but may be replaced in the proof by a fixed point $\xi_0^N \in \{0,1\}^{S_N}$ and so holds equally well in our setting. The derivation of (7.65) is the same. One just replaces $\sigma_x^N(a_0, a_1, a_2)$ with $\sigma_x^N(a_1, a_2)$. Note that (1.7) is used in the proofs with $n = 3$ and $n = 2$, respectively. □

For nonempty sets $A \subset \bar{\mathcal{N}}$ define

$$\begin{aligned}\Theta_2^N(A) &= (\log N)\widehat{P}(\sigma^N(A/\sqrt{N}, (\bar{\mathcal{N}} \setminus A)/\sqrt{N}) > t_N,\ \tau^N(A/\sqrt{N}, (\bar{\mathcal{N}} \setminus A)/\sqrt{N}) < t_N) \\ &= (\log N)\widehat{P}(\sigma(A, \bar{\mathcal{N}} \setminus A) > N w_N t_N,\ \tau(A, \bar{\mathcal{N}} \setminus A) < N w_N t_N) \\ \Theta_2^N &= \sum_{\emptyset \neq A \subset \mathcal{N}} r^{N,a}(A)\Theta_2^N(A).\end{aligned}$$

By Proposition 1.6 and (3.28),

$$\lim_{N\to\infty} \Theta_2^N = \Theta_2. \tag{7.66}$$

Proposition 7.10. *There is a $C_{7.67} > 0$ such that for any $T > 0$, $\phi \in C_b([0,T] \times \mathbb{R}^2)$, and $0 \le s, u \le T$ with $|u - s| \le t_N$,*

$$\Big|\hat{H}^{N,2}(\xi_0^N, t_N, \Phi_s) - \Theta_2^N X_0^N(\Phi_u)\Big| \le C_{7.67}\|\Phi\|\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-1/2} X_0^N(\mathbf{1})\Big). \tag{7.67}$$

Proof. In view of Lemma 7.4 we will start with the difference $\hat{H}_2^{N,2}(\xi_0^N, t_N, \Phi_s) - \Theta_2^N X_0^N(\Phi_s)$. Let $\emptyset \neq A \subset \mathcal{N}_N$, and recall the definition of $\Omega_x^N(A)$ in (7.36), taking $u_N = t_N$. For any fixed $a \in A$,

$$I_2^{N,+}(x, t_N, A, \xi_0^N) = \big[\xi_0^N(B_{t_N}^{N,x+a}) - \xi_0^N(B_{t_N}^{N,x+a})\xi_0^N(B_{t_N}^{N,x})\big]1\{\Omega_x^N(A)\}, \tag{7.68}$$

and thus

$$\begin{aligned}&|\hat{E}\big(I_2^{N,+}(x, t_N, A, \xi_0^N)\big) - \hat{P}\big(B_{t_N}^{N,x+a} \in \xi_0^N, \Omega_x^N(A)\big)| \\ &\qquad \le \hat{P}\big(B_{t_N}^{N,x+a} \in \xi_0^N, B_{t_N}^{N,x} \in \xi_0^N, \sigma^N(x+a, x) > t_N\big).\end{aligned} \tag{7.69}$$

If we choose an $a = a(A) \in A$ for each non-empty $A \subset \mathcal{N}_N$, define

$$\tilde{H}_2^{N,2}(\xi_0^N, t_N, \Phi_s) = \sum_{\emptyset \neq A \subset \mathcal{N}_N} r^{N,a}(\sqrt{N}A)\frac{\log N}{N'} \sum_{x \in S_N} \Phi(s,x)\hat{P}(B_{t_N}^{N,x+a} \in \xi_0^N, \Omega_x^N(A)). \tag{7.70}$$

It then follows from Proposition 7.9 that for a constant C

$$|\hat{H}_2^{N,2}(\xi_0^N, t_N, \Phi_s) - \tilde{H}_2^{N,2}(\xi_0^N, t_N, \Phi_s)| \le C\|\Phi\|_\infty\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-(1/2)} X_0^N(\mathbf{1})\Big). \tag{7.71}$$

To evaluate $\tilde H^{N,2}_{t_N}$, for fixed A we have

$$\begin{aligned}
&\frac{\log N}{N'}\sum_{x\in S_N}\Phi(s,x)\hat P(B^{N,x+a}\in\xi_0^N,\Omega_x^N(A))\\
&\qquad=\frac{\log N}{N'}\sum_{x\in S_N}\Phi(s,x)\sum_{w\in S_N}\xi_0^N(x+w)\hat P(B^{N,x+a}_{t_N}=x+w,\Omega_x^N(A))\\
&\qquad=\frac{\log N}{N'}\sum_{x\in S_N}\sum_{w\in S_N}[\Phi(s,x)-\Phi(s,x+w)]\xi_0^N(x+w)\hat P(B^{N,a}_{t_N}=w,\Omega_0^N(A))\\
&\qquad\quad+\frac{\log N}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Phi(s,x+w)\xi_0^N(x+w)\hat P(B^{N,a}_{t_N}=w,\Omega_0^N(A)).
\end{aligned}$$

As in the proof of (7.33) (see (7.37)),

$$\begin{aligned}
&\frac{\log N}{N'}\sum_{x\in S_N}\sum_{w\in S_N}|\Phi(s,x)-\Phi(s,x+w)|\xi_0^N(x+w)\hat P(B^{N,a}_{t_N}=w)\\
&\qquad\qquad\le C(\mathcal N)\|\Phi_s\|_{\mathtt{Lip}}X_0^N(\mathbf 1)(\log N)^{-17/2}. \qquad (7.72)
\end{aligned}$$

Summing first over x and then over w,

$$\frac{\log N}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Phi(s,x+w)\xi_0^N(x+w)\hat P(B^{N,0}_{t_N}=w,\Omega_0^N(A))=X_0^N(\Phi_s)(\log N)\hat P(\Omega_0^N(A)). \tag{7.73}$$

Combining (7.71)–(7.73), summing over $\emptyset\neq A\subset\mathcal N$, and using the definition of $\Theta^{N,2}$ we get

$$\begin{aligned}
&|\hat H_2^{N,2}(\xi_0^N,t_N,\Phi_s)-\Theta_2^NX_0^N(\Phi_s)|\\
&\qquad\le C\|\Phi\|\Big(\frac{1}{t_N\log N}\mathscr I^N(\xi_0^N)+(\log N)^{-(1/2)}X_0^N(\mathbf 1)\Big).
\end{aligned}\tag{7.74}$$

Now combine this with the easy bound $|X_0^N(\Phi(s,\cdot)-X_0^N(\Phi(u,\cdot)|\le|\Phi|_{1/2,N}\sqrt{t_N}X_0^N(\mathbf 1)$ and the boundedness of $\{\Theta_2^N\}$ to derive (7.67), but with $\hat H_2^{N,2}$ in place of $\hat H^{N,2}$. Finally, use (7.31) for $i\ge 3$ and the above bound in the representation (7.30) for $\hat H^{N,2}$ to complete the proof. □

The proof of Proposition 7.11 below, the analogue of Proposition 7.10 for $\hat H^{N,3}$, is more involved and requires additional notation. For nonempty finite disjoint sets $A,A_1,A_2\subset\bar{\mathcal N}$ let

$$\begin{aligned}
\Theta_3^N&(A,A_1,A_2)\\
&=(\log N)^3\widehat P(\sigma^N(A/\sqrt N,A_1/\sqrt N,A_2/\sqrt N)>t_N,\,\tau^N(A/\sqrt N,A_1/\sqrt N,A_2/\sqrt N)<t_N)\\
&=(\log N)^3\widehat P(\sigma(A,A_1,A_2)>Nw_Nt_N,\,\tau(A,A_1,A_2)<Nw_Nt_N),
\end{aligned}$$

and define

$$\begin{aligned}
\Theta_3^{N,+}(A)&=\sum_{\{A_1,A_2\}\in\mathcal P(\bar{\mathcal N}\setminus A)}\Theta_3^N(A,A_1,A_2),\\
\Theta_3^{N,-}(A)&=\sum_{\{A_1,A_2\}\in\mathcal P(A)}\Theta_3^N(\bar{\mathcal N}\setminus A,A_1,A_2),\\
\Theta_3^{N\pm}&=\sum_{\emptyset\neq A\subset\mathcal N}r^{N,s}(A)\Theta^{N,\pm}(A)
\end{aligned}\tag{7.75}$$

and $\Theta_3^N=\Theta_3^{N,+}-\Theta_3^{N,-}$. Proposition 1.6 and (3.30) imply that if Θ_3 is as in (3.33), then

$$\lim_N\Theta_3^N=\Theta_3. \tag{7.76}$$

Proposition 7.11. *There is a constant $C_{7.77} > 0$ such that for any $T > 0$, $\phi \in C_b([0,T] \times \mathbb{R}^2)$, and $0 \le s, u \le T$ with $|u-s| \le t_N$,*

$$\left|\hat{H}^{N,3}(\xi_0^N, t_N, \Phi_s) - \Theta_3^N X_0^N(\Phi_u)\right| \le C_{7.77}\|\Phi\|\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-1/2} X_0^N(\mathbf{1})\Big). \tag{7.77}$$

Proof. Let us start with the difference $\hat{H}_3^N(\xi_0^N, t_N, \Phi_s) - \Theta^N X_0^N(\Phi(u,\cdot))$. Assume $\emptyset \ne A \subset \mathcal{N}_N$. It is easy to see that if $I_3^N(x, t_N, A, \xi_0^N) \ne 0$, then exactly one of $|B_{t_N}^{N,A}|$ or $|B_{t_N}^{N,\bar{\mathcal{N}}_N \setminus A}|$ must be 2, while the other must be 1. To account for these possibilities, we introduce

$$\begin{aligned}
\chi_1^{N,+}(x, t_N, A, \xi_0^N) &= 1\{B_{t_N}^{N,x+A} \subset \xi_0^N, |B_{t_N}^{N,x+A}| = 2, B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A} \subset \widehat{\xi}_0^N, |B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}| = 1\},\\
\chi_1^{N,-}(x, t_N, A, \xi_0^N) &= 1\{B_{t_N}^{N,x+A} \subset \widehat{\xi}_0^N, |B_{t_N}^{N,x+A}| = 1, B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A} \subset \xi_0^N, |B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}| = 2\},\\
\chi_2^{N,+}(x, t_N, A, \xi_0^N) &= 1\{B_{t_N}^{N,x+A} \subset \xi_0^N, |B_{t_N}^{N,x+A}| = 1, B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A} \subset \widehat{\xi}_0^N, |B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}| = 2\},\\
\chi_2^{N,-}(x, t_N, A, \xi_0^N) &= 1\{B_{t_N}^{N,x+A} \subset \widehat{\xi}_0^N, |B_{t_N}^{N,x+A}| = 2, B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A} \subset \xi_0^N, |B_{t_N}^{N,x+\bar{\mathcal{N}}_N\setminus A}| = 1\},
\end{aligned}$$

and for $i = 1, 2$,

$$\begin{aligned}
&\hat{\Sigma}_3^{N,i}(\xi_0^N, t_N, \Phi_s)\\
&\quad = \sum_{\emptyset \ne A \subsetneq \mathcal{N}_N} r^{N,s}(\sqrt{N}A)\frac{(\log N)^3}{N'}\sum_{x \in S_N}\Phi(s,x)\widehat{E}[\chi_i^{N,+}(x, t_N, A, \xi_0^N) - \chi_i^{N,-}(x, t_N, A, \xi_0^N)].
\end{aligned}$$

It follows that

$$(I_3^{N,+} - I_3^{N,-})(x, t_N, A, \xi_0^N) = (\chi_1^{N,+} - \chi_1^{N,-})(x, t_N, A, \xi_0^N) + (\chi_2^{N,+} - \chi_2^{N,-})(x, t_N, A, \xi_0^N),$$

and thus

$$\hat{H}_3^N(\xi_0^N, t_N, \Phi_s) = \hat{\Sigma}_3^{N,1}(\xi_0^N, t_N, \Phi_s) + \hat{\Sigma}_3^{N,2}(\xi_0^N, t_N, \Phi_s). \tag{7.78}$$

The term $\hat{\Sigma}_3^{N,1}(\xi_0^N, t_N, \Phi_s)$ is small. This is because both $\chi_1^{N,+}$ and $\chi_1^{N,-}$ are bounded above by

$$\sum_{\text{distinct } a_0, a_1, a_2 \in \bar{\mathcal{N}}} 1\{B_{t_N}^{N,x+a_1} \in \xi_0^N, B_{t_N}^{N,x+a_2} \in \xi_0^N, \sigma_x^N(a_0, a_1, a_2) > t_N\},$$

which implies by Lemma 7.9 that

$$|\hat{\Sigma}_3^{N,1}(\xi_0^N, t_N, \Phi_s)| \le C_{7.64}\binom{|\bar{\mathcal{N}}|}{3}\|\Phi\|_\infty\Big(\frac{1}{t_N \log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-1/2}X_0^N(\mathbf{1})\Big). \tag{7.79}$$

To handle $\hat{\Sigma}_3^{N,2}$ it is convenient to define

$$\hat{\Sigma}_3^{N,2,\pm}(\xi_0^N, t_N, \Phi_s) = \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A)\frac{(\log N)^3}{N'}\sum_{x \in S_N}\Phi(s,x)\widehat{E}\big(\chi_2^{N,\pm}(x, t_N, A, \xi_0^N)\big),$$

so that

$$\hat{\Sigma}_3^{N,2}(\xi_0^N, t_N, \Phi_s) = \hat{\Sigma}_3^{N,2,+}(\xi_0^N, t_N, \Phi_s) - \hat{\Sigma}_3^{N,2,-}(\xi_0^N, t_N, \Phi_s). \tag{7.80}$$

Consider $\hat{\Sigma}_3^{N,2,+}$, and let $\emptyset \ne A \subset \mathcal{N}_N$. On the event defining $\chi_2^{N,+}(x, t_N, A, \xi_0^N)$ there must exist nonempty disjoint A_1, A_2 whose union is $\bar{\mathcal{N}}_N \setminus A$, such that none of the walks starting from the three disjoint sets A, A_1 and A_2 meet by time t_N, $|B_{t_N}^{N,x+A_1}| = |B_{t_N}^{N,x+A_2}| = 1$ and $B_{t_N}^{N,x+A_1}, B_{t_N}^{N,x+A_2} \subset \widehat{\xi}_0^N$. Thus, if we define

$$\Omega_x^N(A, A_1, A_2) = \{\sigma_x^N(A, A_1, A_2) > t_N, \tau_x^N(A, A_1, A_2) < t_N\},$$

then for any $a \in A$ and $a_i \in A_i$,

$$\widehat{E}\big(\chi_2^{N,+}(x,t_N,A,\xi_0^N)\big) = \sum_{\substack{\{A_1,A_2\}\in\\ \mathcal{P}(\bar{\mathcal{N}}_N\setminus A)}} \widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} \in \xi_0^N, B_{t_N}^{N,x+a_1} \in \widehat{\xi}_0^N, B_{t_N}^{N,x+a_2} \in \widehat{\xi}_0^N\Big). \tag{7.81}$$

The next step is to see that we may drop requirement that $B_{t_N}^{N,x+a_1}, B^{N,x+a_2} \in \widehat{\xi}_0^N$ above. Observe that

$$\begin{aligned}
&\Big|\widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} \in \xi_0^N, B_{t_N}^{N,x+a_1} \in \widehat{\xi}_0^N, B_{t_N}^{N,x+a_2} \in \widehat{\xi}_0^N\Big)\\
&\qquad\qquad - \widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} \in \xi_0^N\Big)\Big|\\
&\quad \le \widehat{P}\Big(\sigma_x^N(A,A_1,A_2) > t_N, B_{t_N}^{N,x+a} \in \xi_0^N, B_{t_N}^{N,x+a_1} \text{ or } B_{t_N}^{N,x+a_2} \in \xi_0^N\Big)\\
&\quad \le 2 \max_{\text{distinct } a,a_1,a_2\in\bar{\mathcal{N}}_N} \widehat{P}(\sigma_x^N(a,a_1,a_2) > t_N, B_{t_N}^{N,x+a} \in \xi_0^N, B_{t_N}^{N,x+a_1} \in \xi_0^N).
\end{aligned} \tag{7.82}$$

If we let

$$\begin{aligned}
&\widetilde{\Sigma}_3^{N,2,+}(\xi_0^N,t_N,\Phi_s)\\
&\quad = \sum_{A\subset\mathcal{N}} r^{N,s}(\sqrt{N}A)\frac{(\log N)^3}{N'}\sum_{x\in S_N}\Phi(s,x)\sum_{\substack{\{A_1,A_2\}\in\\ \mathcal{P}(\bar{\mathcal{N}}_N\setminus A)}} \widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} \in \xi_0^N\Big),
\end{aligned} \tag{7.83}$$

it follows from the above and Lemma 7.9 that there is a constant $C_{7.84}$ such that

$$\begin{aligned}
&\Big|\hat{\Sigma}_3^{N,2,+}(\xi_0^N,t_N,\Phi_s) - \widetilde{\Sigma}_3^{N,2,+}(\xi_0^N,t_N,\Phi_s)\Big|\\
&\qquad\qquad \le C_{7.84}\|\Phi\|_\infty\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-1/2}X_0^N(\mathbf{1})\Big).
\end{aligned} \tag{7.84}$$

Again, consider a fixed $\emptyset \ne A \subset \mathcal{N}_N$ and let $\{A_1,A_2\} \in \mathcal{P}(\bar{\mathcal{N}}_N \setminus A)$. By translation invariance,

$$\begin{aligned}
&\frac{(\log N)^3}{N'}\sum_{x\in S_N}\Phi(s,x)\widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} \in \xi_0^N)\\
&\quad = \frac{(\log N)^3}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Phi(s,x)\xi_0^N(x+w)\widehat{P}\Big(\Omega_x^N(A,A_1,A_2), B_{t_N}^{N,x+a} = x+w\Big)\\
&\quad = \frac{(\log N)^3}{N'}\sum_{x\in S_N}\sum_{w\in S_N}[\Phi(s,x) - \Phi(s,x+w)]\xi_0^N(x+w)\widehat{P}\Big(\Omega_0^N(A,A_1,A_2), B_{t_N}^{N,a} = w\Big)\\
&\qquad + \frac{(\log N)^3}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Phi(s,x+w)\xi_0^N(x+w)\widehat{P}\Big(\Omega_0^N(A,A_1,A_2), B_{t_N}^{N,a} = w\Big).
\end{aligned} \tag{7.85}$$

As in the proof of (7.33) (see (7.37)),

$$\begin{aligned}
&\frac{(\log N)^3}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Big|(\Phi(s,x) - \Phi(s,x+w))\Big|\xi_0^N(x+w)\widehat{P}\Big(\Omega_0^N(A,A_1,A_2), B_{t_N}^{N,a} = w\Big)\\
&\qquad\qquad \le C_{7.86}\|\Phi\|_{\mathtt{Lip}}X_0^N(\mathbf{1})(\log N)^{-13/2}.
\end{aligned} \tag{7.86}$$

For the final sum in (7.85), summing first over x and then over w yields

$$\frac{(\log N)^3}{N'}\sum_{x\in S_N}\sum_{w\in S_N}\Phi(s,x+w)\xi_0^N(x+w)\widehat{P}\Big(\Omega_0^N(A,A_1,A_2),B_{t_N}^{N,a}=w\Big)$$
$$=\Theta_3^N(\sqrt{N}A,\sqrt{N}A_1,\sqrt{N}A_2)X_0^N(\Phi_s). \quad (7.87)$$

Combining (7.85)–(7.87) and using the definition of $\tilde\Sigma$ we find that

$$\Big|\frac{(\log N)^3}{N'}\sum_{x\in S_N}\Phi(s,x)\widehat{P}\Big(\Omega_x^N(A,A_1,A_2),B_{t_N}^{N,x+a}\in\xi_0^N\Big)-\Theta_3^N(\sqrt{N}A,\sqrt{N}A_1,\sqrt{N}A_2)X_0^N(\Phi_s)\Big|$$
$$\le C_{7.86}\|\Phi\|_{\mathtt{Lip}}X_0^N(\mathbf{1})(\log N)^{-13/2}, \quad (7.88)$$

$$\Big|\tilde\Sigma^{N,2,+}(\xi_0^N,t_N,\Phi_s)-\Theta^{N,+}X_0^N(\Phi_s)\Big|\le C_{7.86}\binom{|\bar{\mathcal{N}}_N|}{3}(\log N)^{-13/2}\|\Phi\|_{\mathtt{Lip}}X_0^N(\mathbf{1}), \quad (7.89)$$

and by (7.84),

$$\Big|\hat\Sigma_3^{N,2,+}(\xi_0^N,t_N,\Phi_s)-\Theta^{N,+}X_0^N(\Phi_s)\Big|$$
$$\le C_{7.90}\|\Phi\|_{\mathtt{Lip}}\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_0^N)+(\log N)^{-1/2}X_0^N(\mathbf{1})\Big). \quad (7.90)$$

In the above we have rescaled the sets and summed over subsets of $\mathcal{N}$ (or $\bar{\mathcal{N}}$) instead of $\mathcal{N}_N$ (or $\bar{\mathcal{N}}_N$). The same bound holds for the difference $\hat\Sigma_3^{N,2,-}(\xi_0^N,t_N,\Phi_s)-\Theta^{N,-}X_0^N(\Phi_s)$, and thus in view of (7.78), (7.79), (7.80), (7.90), and the simple bound $|X_0^N(\Phi(s,\cdot)-X_0^N(\Phi(u,\cdot)|\le|\Phi|_{1/2,N}\sqrt{t_N}X_0^N(\mathbf{1})$, we conclude that

$$\Big|\hat H_3^N(\xi_0^N,t_N,\Phi_s)-\Theta^N X_0^N(\Phi_u)\Big|\le C_{7.77}\|\Phi\|\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_0^N)+(\log N)^{-1/2}X_0^N(\mathbf{1})\Big).$$

Combine this with Lemma 7.4 and (7.30) to obtain (7.77). □

As an immediate consequence of Lemma 7.3 (with $u_N=t_N$), Proposition 7.10 and Proposition 7.11 we obtain the following result.

Proposition 7.12. *There is a constant $C_{7.91}>0$ such that for $j=2,3$, any $T>0$, $\Phi\in C_b([0,T]\times\mathbb{R}^2)$, and $s\in[t_N,T]$,*

$$\Big|E\big(d^{N,j}(s,\xi_s^N,\Phi)|\mathcal{F}_{s-t_N}^N\big)-\Theta_j^N X_{s-t_N}^N(\Phi)\Big|$$
$$\le C_{7.91}\|\Phi\|\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_{s-t_N}^N)+(\log N)^{-1/2}X_{s-t_N}^N(\mathbf{1})\Big). \quad (7.91)$$

Turning briefly to the martingale square function, and recalling (7.50) and the definition of K_N from (7.47), we have the following analogue of the above.

Proposition 7.13. *There is a $C_{7.92}$ such that for any bounded Lipschitz continuous $\Phi:\mathbb{R}^2\to\mathbb{R}$ and $s\ge t_N$,*

$$\Big|E\big(X_s^N(\log N\Phi f_0^{(N)}(\cdot,\xi_s^N))|\mathcal{F}_{s-t_N}^N\big)-K_N X_{s-t_N}^N(\Phi)\Big|$$
$$\le C_{7.92}\|\Phi\|_{Lip}\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_{s-t_N}^N)+(\log N)^{-1/2}X_{s-t_N}^N(\mathbf{1})\Big). \quad (7.92)$$

Proof. By (7.56) and (7.57),

$$E(X^N_{t_N}(\log N\Phi f_0^{(N)}(\cdot,\xi^N_{t_N})) = \frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}_N}\Phi(x)p_N(y)\hat{E}(\xi_0^N(B^{N,x}_{t_N})\widehat{\xi}_0^N(B^{N,x+y}_{t_N})) + R_N \tag{7.93}$$

$$:= S_N + R_N,$$

where

$$|R_N| \le C\|\Phi\|_\infty(\log N)^{-15}X_0^N(\mathbf{1}). \tag{7.94}$$

Clearly

$$\begin{aligned} S_N &= \frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}_N}\Phi(x)p_N(y)\hat{E}(\xi_0^N(B^{N,x}_{t_N})1(\sigma^N(x,x+y)>t_N)) \\ &\quad - \frac{\log N}{N'}\sum_{x\in S_N}\sum_{y\in\mathcal{N}_N}\Phi(x)p_N(y)\hat{E}(\xi_0^N(B^{N,x}_{t_N})\xi_0^N(B^{N,x+y}_{t_N})1(\sigma^N(x,x+y)>t_N) \\ &:= S_N^1 - S_N^2. \end{aligned}$$

(7.65) shows that

$$|S_N^2| \le C_{7.64}\|\Phi\|_\infty\Big(\frac{1}{t_N\log N}\mathscr{I}^N(\xi_0^N) + (\log N)^{-(1/2)}X_0^N(\mathbf{1})\Big). \tag{7.95}$$

By Lemma 7.8 (a),

$$S_N^1 = K_N X_0^N(\Phi) + \mathcal{E}_N,$$

where

$$|\mathcal{E}_N| \le C_{7.54}|\Phi|_{\mathtt{Lip}}(\log N)^{-17/2}X_0^N(\mathbf{1}). \tag{7.96}$$

Use the error bounds (7.94), (7.95) and (7.96) in the decomposition (7.93), and then apply the Markov property to complete the proof. □

In view of the above results, to obtain exact asymptotics for the limiting drift arising from the q-voter perturbation term as well as the square function of the martingale term, we will show that $\frac{1}{t_N\log N}E\left[\int_{t_N}^{t\vee t_N}\mathscr{I}^N(\xi^N_{s-t_N})ds\right]$ is negligible for large N.

Proposition 7.14. *Let $(\delta_N)_{N\ge3}$ be a positive sequence such that $\lim_{N\to\infty}\delta_N = 0$ and $\liminf_{N\to\infty}\sqrt{N}\delta_N > 0$. For any T, there exists $C_{7.97}$, depending on T and $\{\delta_N\}$, such that for any $t\le T$, for any $N\ge3$,*

$$E[\mathscr{I}^N(\delta_N,\xi^N_t)] \le C_{7.97}\Big(X_0^N(1) + X_0^N(1)^2\Big)\Big[\delta_N\Big(1+\log\Big(1+\frac{t}{\delta_N}\Big)\Big) + \frac{\delta_N}{t+\delta_N}\Big]. \tag{7.97}$$

The proof of this key result is similar in outline to that of the same result, Proposition 3.10 in [9], for Lotka-Volterra models, although additional complications will arise in the present setting. We will prove it in Section 9 below but assume it in what follows.

Corollary 7.15. *If $T>0$ there exists a constant $C_{7.98}$ depending on T such that for any $t\le T$,*

$$\frac{1}{t_N\log N}E\left[\int_{t_N}^{t\vee t_N}\mathscr{I}^N(\xi^N_{s-t_N})ds\right] \le C_{7.98}(\log N)^{-1/2}(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2). \tag{7.98}$$

Proof. Given Proposition 7.14, the elementary proof of the above is identical to that of Corollary 3.11 in [9]. (The bound in the final line of the proof there leads to an upper bound as above but instead with $(\log N)^{-\eta+\delta}$ for any $\delta>0$ and $\eta\in(0,1)$, and hence, in particular, $(\log N)^{-1/2}$). □

We are finally ready to give the asymptotic behavior of the drifts, $D_t^{N,j}(\Phi)$, $j=2,3$, and so prove the main result of this Section.

Proposition 7.16. *For $j=2,3$, any $\Phi:\mathbb{R}^2\to\mathbb{R}$ with $\|\Phi\|_{Lip}<\infty$ and all $t>0$,*

$$\lim_{N\to\infty} E\Big(\Big|D_t^{N,j}(\Phi)-\Theta_j\int_0^t X_s^N(\Phi)\,ds\Big|\Big)=0.$$

Proof. Let $t\in(0,T]$. A little thought shows the above expectation is bounded by

$$\begin{aligned}
&E\Big(\int_0^{t_N\wedge t}|d^{N,j}(s,\xi_s^N,\Phi)|\,ds\Big)+E\Big(\Big(\int_{t_N}^{t_N\vee t}d^{N,j}(s,\xi_s^N,\Phi)-E(d^{N,j}(s,\xi_s^N,\Phi)|\mathcal{F}_{s-t_N}^N)\,ds\Big)^2\Big)^{1/2}\\
&\quad+E\Big(\int_{t_N}^{t_N\vee t}|E(d^{N,j}(s,\xi_s^N,\Phi)|\mathcal{F}_{s-t_N}^N)-\Theta_j^N X_{s-t_N}^N(\Phi)|\,ds\Big) \qquad (7.99)\\
&\quad+|\Theta_j^N-\Theta_j|\|\Phi\|_\infty E\Big(\int_0^{(t-t_N)^+}X_s^N(\mathbf{1})\,ds\Big)+|\Theta_j|\|\Phi\|_\infty E\Big(\int_{(t-t_N)^+}^t X_s^N(\mathbf{1})\,ds\Big).
\end{aligned}$$

Proposition 7.6(a) and (7.39) shows the first term is at most

$$C\|\Phi\|_\infty(\log N)^3 t_N X_0^N(\mathbf{1})\to 0 \text{ as } N\to\infty.$$

Use an orthogonality argument as in the derivation of (67) in [9] and also (7.39) to bound the second term in (7.99) by

$$C\|\Phi\|_\infty(\log N)^3 t_N^{1/2}\Big[E\Big(\int_{t_N}^{t_N\vee t}X_s^N(\mathbf{1})^2\,ds\Big)\Big]^{1/2}\le C_T\|\Phi\|_\infty(\log N)^{-13/2}(1+X_0^N(\mathbf{1}))\to 0 \text{ as } N\to\infty,$$

where the inequality holds by Proposition 7.6(b). Since $|\Theta_j^N-\Theta_j|\to 0$ (by (7.66) and (7.76)), we may apply Proposition 7.6(a) to see that the last two terms in (7.99) are bounded by

$$C_T\|\Phi\|_\infty X_0^N(\mathbf{1})[|\Theta_j^N-\Theta_j|+t_N]\to 0 \text{ as } N\to\infty.$$

Finally we may use Proposition 7.12, Corollary 7.15 and Proposition 7.6(a) to bound the remaining (middle) term of (7.99) by

$$C_T\|\Phi\|_{\mathtt{Lip}}(\log N)^{-1/2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)\to 0 \text{ as } N\to\infty.$$

Combining the above displays, we may complete the proof. □

Remark 7.17. If we keep track of the bounds in the above proof (and drop the $|\Theta_j^N-\Theta_j|$ term) we get for $j=2,3$,

$$E\Big(\Big|D_t^{N,j}(\Phi)-\Theta_j^N\int_0^t X_s^N(\Phi)\,ds\Big|\Big)\le C_T\|\Phi\|_{\mathtt{Lip}}(1+X_0^N(\mathbf{1})^2)(\log N)^{-1/2}\quad\text{for all } t\in[0,T]. \qquad (7.100)$$

Proposition 7.18. *For any $\Phi:\mathbb{R}^2\to\mathbb{R}$ with $\|\Phi\|_{Lip}<\infty$ and all $t>0$,*

$$\lim_{N\to\infty}E\Big(\Big|\langle M^N(\Phi)\rangle_t-4\pi\sigma^2\int_0^t X_s^N(\Phi^2)\,ds\Big|\Big)=0.$$

Proof. Note that $\|\Phi\|_{\mathtt{Lip}}$ finite implies that $\|\Phi^2\|_{\mathtt{Lip}}$ is finite. Arguing exactly as in the proof of Proposition 7.16, using Proposition 7.13 in place of Proposition 7.12 and (7.48) in place of (7.66) and (7.76), we get

$$\lim_{N\to\infty}E\Big(\Big|\int_0^t X_s^N(2\log N\Phi^2 f_0^{(N)}(\cdot,\xi_s^N))\,ds-\int_0^t 4\pi\sigma^2X_s^N(\Phi^2)\,ds\Big|\Big)=0.$$

Now use (7.50), (7.51), and Proposition 7.6(a) to complete the proof. □

Define $(P_t^N, t \geq 0)$ as the semigroup of the rate-N random walk on S_N with jump kernel p_N. By translation invariance we can have P_t^N operate on functions on the plane, even though S_N is the natural state space.

As an application of the control of the drift terms $d^{N,j}$ given by Proposition 7.16, we obtain an effective bound on the mean measures of our rescaled q-voter models.

Lemma 7.19. *There exists a $c_{7.101} > 0$, and for any $T > 0$ a constant $C_{7.101}(T) > 0$ so that for all $t \in [0,T]$ and any $\Psi : \mathbb{R}^2 \to \mathbb{R}_+$ such that $\|\Psi\|_{\mathtt{Lip}} \leq T$,*

$$E[X_t^N(\Psi)] \leq e^{c_{7.101}t} X_0^N(P_t^N(\Psi)) + C_{7.101}(\log N)^{-1/2}(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2). \tag{7.101}$$

Proof. Fix $T > 0$, let $t \in [0,T]$ and Ψ be as in the Lemma, and let $c \in \mathbb{R}$. Define $\Phi \in C_b([0,t] \times \mathbb{R}^d)$ by $\Phi(s,x) = e^{-cs} P_{t-s}^N \Psi(x)$, so that

$$\dot{\Phi}(s,x) = -A_N \Phi(s,x) - c\Phi(s,x) \in C_b([0,t] \times \mathbb{R}^2). \tag{7.102}$$

Let $d^N(s, \xi_s^N, \Phi) = d^{N,2}(s, \xi_s^N, \Phi) + d^{N,3}(s, \xi_s^N, \Phi)$ and $\Theta^N = \Theta_2^N + \Theta_3^N$. By (6.10) and (7.102) we have

$$E(X_t^N(\Phi(t,\cdot))) = X_0^N(P_t^N(\Psi)) + \int_0^t E(d^N(s, \xi_s^N, \Phi) - cX_s^N(\Phi))\, ds. \tag{7.103}$$

Now choose $c > 1 \vee \sigma \vee (\sup_{N \geq e^3} \Theta^N)$ (recall (7.66) and (7.76)). It is then easy to see (and in fact is shown in the proof of Lemma 3.5 in [9]–see p. 1232) that $c > 1 \vee \sigma$ implies

$$\|\Phi\| \leq 2c\|\Psi\|_{\mathtt{Lip}} \leq 2cT, \tag{7.104}$$

where the last inequality is by hypothesis. Now use (7.100) and (7.104) to deduce that

$$E\Big(\int_0^t d^N(s, \xi_s^N, \Phi)\, ds\Big) \leq \int_0^t cE(X_s^N(\Phi))\, ds + C_T\|\Psi\|_{\mathtt{Lip}}(\log N)^{-1/2}(1 + X_0^N(\mathbf{1})^2), \;\; \forall t \in [0,T].$$

Finally use this in (7.103), noting that $\Phi(t,x) = e^{-ct}\Psi(x)$, to complete the proof with $c_{7.101} = c$. ☐

8 Convergence to super-Brownian motion: Proof of Theorem 6.1

We assume the hypotheses of Theorem 6.1 whose proof is the objective of this section.

8.1 Relative compactness

A collection of stochastic processes $\{Y^N : N \geq \alpha\}$ with sample paths in $D(\mathbb{R}_+, S)$ for some Polish space S is C-relatively compact iff for every sequence $N_k \uparrow \infty$ in $[\alpha, \infty)$ there is a subsequence $\{N_k'\}$ so that $Y^{N_k'}$ converges weakly in $D(\mathbb{R}_+, S)$ to a process with continuous paths. The same definition applies to a given *sequence* of processes.

Proposition 8.1. *The set $\{X^N, N \geq N(\varepsilon_0)\}$ is C-relatively compact in $D(\mathbb{R}^+, \mathcal{M}_F(\mathbb{R}^2))$.*

To prove this it clearly suffices to show that for every $N_k \uparrow \infty$ $(N_k \geq N(\varepsilon_0))$, $\{X^{N_k}\}$, is C-relatively compact. To ease the notation we take $N_k = k$ as the proof in the general case is the same. Hence we reduce to the case of showing $\{X^N : N \in \mathbb{N}, N \geq N(\varepsilon_0)\}$ is relatively compact. This result will follow from Jakubowski's theorem (see Theorem II.4.1 in [P2002]) and the following two lemmas.

Lemma 8.2. *For any $\Phi \in C_b^3(\mathbb{R}^2)$, the sequence $\{X^N(\Phi), N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$ is C-relatively compact in $D(\mathbb{R}_+, \mathbb{R})$.*

Lemma 8.3. *For any $\varepsilon > 0$, $T > 0$ there exists $A > 0$ such that*

$$\sup_{N \in \mathbb{N}^{\geq N(\varepsilon_0)}} P\left(\sup_{t \leq T} X_t^N(B(0,A)^c) > \varepsilon\right) < \varepsilon.$$

Proof of Lemma 8.2. We use (6.10) to establish C-relative compactness of $\{X^N(\Phi) : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$ by establishing the C-relative compactness of each of the terms on the right-hand side. For the latter we proceed as in Lemma 6.1 of [9]. Consider first $\{D^{N,j}(\Phi) : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$ for $j = 2$ or 3. Use Proposition 7.6(b) in (7.60) and $\min(a,b) \leq \sqrt{ab}$ for $a, b > 0$, to see that for $t_N \leq s_1 < s_2 \leq T$,

$$\begin{aligned} E((D^{N,j}_{s_2}(\Phi) - D^{N,j}_{s_1}(\Phi))^2) &\leq C_T \|\Phi\|^2_{\mathtt{Lip}}[(s_2 - s_1)^2 + \log^6 N\sqrt{t_N}(s_2 - s_1)^{3/2}](X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2) \\ &\leq C_T \|\Phi\|^2_{\mathtt{Lip}}(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^{3/2}. \end{aligned} \tag{8.1}$$

Moreover, we have from (7.39) that for $j = 2, 3$,

$$|d^{N,j}(s, \xi^N_s, \Phi)| \leq |\bar{\mathcal{N}}|(\log N)^3 \|\Phi\|_\infty X^N_s(\mathbf{1}), \tag{8.2}$$

and so by Proposition 7.6(b) for $0 \leq s_1 < s_2 \leq t_N$,

$$\begin{aligned} E\Big(\Big(\int_{s_1}^{s_2} |d^{N,j}(s, \xi^N_s, \Phi)|\,ds\Big)^2\Big) &\leq C|\bar{\mathcal{N}}|\|\Phi\|^2_\infty(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(\log N)^6(s_2 - s_1)^2 \\ &\leq C\|\Phi\|^2_\infty(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(\log N)^6\sqrt{t_N}(s_2 - s_1)^{3/2} \\ &\leq C\|\Phi\|^2_\infty(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^{3/2}. \end{aligned} \tag{8.3}$$

The C-relative compactness of $\{D^{N,j}(\Phi) : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$ is now immediate from $D^{N,j}_0 = 0$, (8.1), (8.3) and Kolmogorov's criterion.

Turning to the C-relative compactness of $\{M^N(\Phi) : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$, as in Lemma 6.1 of [9], because the maximum jump of $M^N(\Phi)$ goes to zero as $N \to \infty$, it suffices to prove C-relative compactness of $\{\langle M^N(\Phi)\rangle : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$. From (6.11), (7.50), (7.51), and the second moment bound in Proposition 7.6(b), we have for $0 \leq s_1 < s_2 \leq T$,

$$\begin{aligned} E\Big(\Big(\langle M^N(\Phi)\rangle_{s_2} - \langle M^N(\Phi)\rangle_{s_1}\Big)^2\Big) &\leq C_T \|\Phi\|^4_{\mathtt{Lip}}(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^2 \\ &\quad + C\|\Phi\|^4_\infty E\Big(\Big(\int_{s_1}^{s_2} Z^N_s\,ds\Big)^2\Big), \end{aligned} \tag{8.4}$$

where

$$Z^N_s = X^N_s(\log N f^{(N)}_0(\xi^N_s)) \leq \log N X^N_s(\mathbf{1}). \tag{8.5}$$

Using (7.54), the above upper bound, and arguing exactly as in the proof of (104)–(106) in Lemma 6.1 of [9] (conditioning back in time by t_N and employing the Markov property) we can show that for $T \geq s_2 > s_1 \geq t_N$,

$$E\Big(\Big(\int_{s_1}^{s_2} Z^N_s\,ds\Big)^2\Big) \leq C_T(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^{3/2}. \tag{8.6}$$

For $0 \leq s_1 < s_2 \leq t_N$ we may argue as in (8.3) using (7.51), (7.50), and the upper bound in (8.5), to see that

$$E\Big(\Big(\langle M^N(\Phi)\rangle_{s_2} - \langle M^N(\Phi)\rangle_{s_1}\Big)^2\Big) \leq C_T \|\Phi\|^4_{\mathtt{Lip}}(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^{3/2}.$$

Combine the above with (8.4), and (8.6) to conclude that

$$E\Big(\Big(\langle M^N(\Phi)\rangle_{s_2} - \langle M^N(\Phi)\rangle_{s_1}\Big)^2\Big) \leq C_T \|\Phi\|^4_{\mathtt{Lip}}(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^{3/2} \text{ for } 0 \leq s_1 < s_2 \leq T. \tag{8.7}$$

The C-relative compactness of $\{\langle M^N(\Phi)\rangle : N \in \mathbb{N}^{\geq N(\varepsilon_0)}\}$ now follows from Kolmogorov's criterion.

Finally, the simple argument in Lemma 6.1 of [9] using Proposition 7.6(b) shows for $0 \le s_1 < s_2 \le T$,

$$E\Big(\Big(D^{N,1}_{s_2}(\Phi) - D^{N,1}_{s_1}(\Phi)\Big)^2\Big) \le C_{T,\Phi}(X^N_0(\mathbf{1}) + X^N_0(\mathbf{1})^2)(s_2 - s_1)^2. \tag{8.8}$$

(This is one place where the assumed regularity of Φ is used.) The C-relative compactness of $\{D^{N,1}(\Phi) : N \in \mathbb{N}^{\ge N(\varepsilon_0)}\}$ follows as usual. As $X^N_0(\Phi) \to X_0(\Phi)$ by hypothesis, the above results give the C-relative compactness of $\{X^N(\Phi)\}$. □

Proof of Lemma 8.3. Let $\{h_n : n \in \mathbb{N}\}$ be a sequence of $[0,1]$-valued functions in $C^3_b(\mathbb{R}^2)$ such that

$$\mathbf{1}_{\{|x|>n+1\}} \le h_n(x) \le \mathbf{1}_{\{|x|>n\}}, \text{ and } \sup_n \|h_n\|_{\mathtt{Lip}} \le C. \tag{8.9}$$

For example, if $h : \mathbb{R} \to [0,1]$ is C^∞, increasing, and $1_{\{z>1\}} \le h(z) \le 1_{\{z>0\}}$ we can take $h_n(x) = h(|x| - n)$. It clearly suffices to show that for each fixed $\varepsilon, T > 0$,

$$\lim_{n\to\infty} \sup_{N\ge N(\varepsilon_0)} P\Big(\sup_{t\le T} X^N_t(h_n) \ge \varepsilon\Big) = 0. \tag{8.10}$$

By (6.10)

$$X^N_t(h_n) = M^N_t(h_n) + Y^N_t(h_n), \tag{8.11}$$

where

$$Y^N_t(h_n) = X^N_0(h_n) + \int_0^t X^N_s(A_N h_n)\,ds + D^{N,2}_t(h_n) + D^{N,3}_t(h_n).$$

Now argue as in the derivation of (110) in the proof of Lemma 6.1 in [9], using (6.11), (7.50), (7.51), the second inequality in (8.9), and (7.54) to conclude that for some $\varepsilon_N \to 0$, independent of n,

$$E(\langle M^N(h_n)\rangle_T) \le \varepsilon_N + \int_0^T C_T E(X^N_s(h_n^2))\,ds \le \varepsilon_N + \int_0^T C_T X^N_0(P^N_s(h_n^2))\,ds. \tag{8.12}$$

In the last inequality we used Lemma 7.19 and absorbed some of the constants and terms there into the C_T and ε_N. Chebychev's inequality (recall (8.9)) shows that for any $K > 0$,

$$\lim_{n\to\infty} \sup_{N\ge N(\varepsilon_0)} \sup_{|x|\le K, s\le T} P^N_s(h_n)(x) = 0.$$

The tightness of $\{X^N_0\}$ now shows

$$\lim_{n\to 0} \sup_{N\ge N(\varepsilon_0)} \sup_{s\le T} X^N_0(P^N_s(h_n)) = 0, \tag{8.13}$$

and so the integral on the righthand side of (8.12) approaches 0 as $n \to \infty$, uniformly in N. For each N fixed it is elementary to use (6.12) to see that $\lim_n E(\langle M^N(h_n)\rangle_T) = 0$. It now follows from Doob's strong inequality L^2 and the above that

$$\lim_{n\to\infty} \sup_{N\ge N(\varepsilon_0)} E(\sup_{t\le T} M^N_t(h_n)^2) = 0. \tag{8.14}$$

To prove (8.10), by (8.11) it now clearly suffices to show

$$\lim_{n\to\infty} \sup_{N\ge N(\varepsilon_0)} P(\sup_{t\le T} |Y^N_t(h_n)| \ge \varepsilon) = 0. \tag{8.15}$$

This is (115) in the proof of Lemma 6.1 in [9] for the Lotka-Volterra model, and the proof given there now goes through without change in our more general setting. The required inputs are (8.1), (8.3), (8.8), (8.14), Lemma 7.19, and (8.13). □

8.2 Identification of the limit

Proof of Theorem 1.11. By the C-relative compactness, established above, it remains only to show that the sequential limit points of $\{X^N\}$ coincide. By the Skorokhod representation theorem we may assume that for a sequence $N_k \uparrow \infty,\ N_k \geq N(\varepsilon_0)$, we have

$$X^{N_k} \to X \in C(R_+, M_F(\mathbb{R}^2)) \quad \text{a.s.}$$

We will take limits in (6.10) and (6.11) to see that the law of X satisfies (MP), the martingale problem characterizing the law of SBM$(X_0, , 4\pi\sigma^2, \sigma^2, \Theta)$. To see this, fix $\Phi \in C_b^3(\mathbb{R}^2)$. By Lemma 2.6 of [6] we have

$$\|A_N\Phi - (\sigma^2/2)\Delta\Phi\|_\infty \to 0 \text{ as } N \to \infty,$$

and therefore by Proposition 7.6(a) for all $t > 0$,

$$\lim_{N\to\infty} E\Big(\Big|\int_0^t X_s^N(A_N\Phi)\,ds - \int_0^t X_s^N((\sigma^2/2)\Delta\Phi)\,ds\Big|\Big) = 0. \tag{8.16}$$

One now can use Propositions 7.16 and 7.18, and (8.16) to argue exactly as in the proof of Proposition 3.2 of [10] and take limits along $\{N_k\}$ in (6.10), (6.12) to see that X satisfies (MP) and so is SBM$(X_0, , 4\pi\sigma^2, \sigma^2, \Theta)$. (Only the last two paragraphs of the proof there are used.) The other inputs needed there are the C-relative compactness of $\{X^N(\Phi)\}$, $\{M^N(\Phi)\}$, and $\{D^{N,j}(\Phi)\}$ for $j = 1, 2, 3$, established in the proof of Lemma 8.2, as well as (take $s_1 = 0, s_2 = T$ in (8.7))

$$E((\langle M^N(\Phi)\rangle_T)^2) \leq C(T, \Phi) \ \text{ for all } N.$$

The latter allows one to conclude that the limiting $M(\phi)$ is a martingale with the appropriate square function. □

9 Proof of Proposition 7.14

We need an elementary bound for $p_t^N(x) := NP(B_t^{N,0} = x)$:

$$p_t^N(x) \leq \frac{C_{9.1}}{t} \text{ for all } t > 0,\ x \in S_N,\ N \geq 3. \tag{9.1}$$

For example, see (A7) in [6].

Assume $\delta_N > 0$ converges to zero, and also satisfies

$$\liminf_N \sqrt{N}\delta_N > 0. \tag{9.2}$$

Fix $T \geq 1$ and consider $0 \leq t \leq T$. Let $p_s^{N,z}(x) = p_s^N(z - x)$ and $\phi_s^z = p_{t-s+\delta_N}^{N,z}$. Argue as in the proof of Proposition 3.10 in Section 5.2 of [9], using the semimartingale decomposition (6.10) to see that for a universal constant $C_{9.3}$ we have

$$E\Big(\int\int 1_{\{|x-y|\leq\sqrt{\delta_N}\}}dX_t^N(x)dX_t^N(y)\Big) \leq C_{9.3}[\mathcal{T}_0 + \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3], \tag{9.3}$$

where

$$\mathcal{T}_0 = \frac{\delta_N}{N}\sum_{z\in S_N} X_0^N(p_{t+\delta_N}^{N,z})^2 = \int\int \delta_N p^N_{2(t+\delta_N)}(y-x)dX_0^N(x)dX_0^N(y),$$

$$\mathcal{T}_1 = \frac{\delta_N}{N}\sum_{z\in S_N} E(\langle M(\phi^z)\rangle_{1,t}),$$

$$\mathcal{T}_2 = \frac{\delta_N}{N}\sum_{z\in S_N} E(\langle M(\phi^z)\rangle_{2,t}),$$

$$\mathcal{T}_3 = \sum_{j=2}^3 \frac{\delta_N}{N}\sum_{z\in S_N} E\Big(\Big(\int_0^t d^{N,j}(s,\xi_s^N,\phi^z)\,ds\Big)^2\Big) := \sum_{j=2}^3 \mathcal{T}_3^{(j)}.$$

By (9.1),

$$\mathcal{T}_0 \le \frac{C_{9.4}\delta_N}{t+\delta_N} X_0^N(\mathbf{1})^2. \tag{9.4}$$

For $\mathcal{T}_2$, use (6.12), $|c^{N,a}(x,\xi_s^N)| = |d^{N,2}(x,\xi_s^N)|$, $|c^{N,s}(x,\xi_s^N)| = |d^{N,3}(x,\xi_s^N)|$, and then Chapman-Kolmogorov and (9.1), to conclude that

$$\begin{aligned}\mathcal{T}_2 &\le \frac{\delta_N}{N} E\Big(\int_0^t \frac{(\log N)^3}{(N')^2}\sum_{x\in S_N}\sum_{z\in S_N} p^N_{t-s+\delta_N}(z-x)^2\Big[|d^{N,2}(x,\xi_s^N)| + |d^{N,3}(x,\xi_s^N)|\Big]\,ds\Big)\\ &\le C\frac{\delta_N}{N}(\log N)^3 E\Big(\int_0^t \frac{N}{(N')^2}(t-s+\delta_N)^{-1}\sum_{x\in S_N}[|d^{N,2}(x,\xi_s^N)| + |d^{N,3}(x,\xi_s^N)|]\,ds\Big).\end{aligned} \tag{9.5}$$

From the bounds in (7.2) we have

$$\frac{1}{N'}\sum_{x\in S_N}|d^{N,j}(x.\xi_s^N)| \le \frac{\|r\|}{N'}\sum_{x\in S_N}\sum_{e\in\bar{\mathcal{N}}_N}\xi_s^N(x+e) = \|r\||\bar{\mathcal{N}}|X_s^N(\mathbf{1}).$$

Use the above in (9.5) and conclude from Proposition 7.6(a) that

$$\mathcal{T}_2 \le \frac{\delta_N(\log N)^4}{N}E\Big(\int_0^t \frac{C}{t-s+\delta_N}X_s^N(\mathbf{1})\,ds\Big) \le C_T\delta_N X_0^N(\mathbf{1})\log\Big(1+\frac{t}{\delta_N}\Big)\frac{(\log N)^4}{N}. \tag{9.6}$$

Turning to $\mathcal{T}_1$, from (6.12) and Chapman-Kolmogorov, we have

$$\begin{aligned}\mathcal{T}_1 &= \delta_N\log N E\Big(\int_0^t \frac{1}{N'}p^N_{2(t-s)+2\delta_N}(0)\Big[\sum_{x,y\in S_N}p_N(y-x)(\xi_s^N(x)\widehat{\xi}_s^N(y)+\widehat{\xi}_s^N(x)\xi_s^N(y))\Big]\,ds\Big)\\ &\le C_{9.7}\delta_N\log N\int_0^t (t-s+\delta_N)^{-1}E\Big(\frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi_s^N(x)\widehat{\xi}_s^N(x+e)\Big)\,ds,\end{aligned} \tag{9.7}$$

the last by (9.1). Set $u_N = \frac{\delta_N}{2}\wedge(\log N)^{-p}$ for some $p\ge 11$. Note that (9.2) shows that u_N satisfies (7.12). If $\xi^N\in\{0,1\}^{S_N}$, define

$$G^N(\xi^N) = \frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi^N(x)\widehat{\xi}^N(x+e),$$

and for $s\ge u_N$, let

$$\Delta_N(s) = \Big|E(G^N(\xi_s^N)|\mathcal{F}^N_{s-u_N}) - \widehat{E}\Big(\frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi^N_{s-u_N}(B^{N,x}_{u_N})\hat{\xi}^N_{s-u_n}(B^{N,x+e}_{u_N})\Big)\Big|. \tag{9.8}$$

The Markov property and coalescing duality for the voter model show that

$$\Delta_N(s) = \Big|E_{\xi^N_{s-u_N}}(G^N(\xi^N_{u_N})) - E_{\xi^N_{s-u_N}}(G^N(\xi^{N,\mathsf{vm}}_{u_N}))\Big|. \tag{9.9}$$

We assume that ξ^N, $\overline{\xi}^N$ (the biased voter model with rates as in (7.6)), and $\xi^{N,\mathsf{vm}}$ are constructed as in (6.2) and (7.11), all starting at ξ^0_N (with finitely many ones as usual). Now use the elementary inequality

$$|\prod_{i=1}^m \xi^N(x_i)\prod_{i=m+1}^{m+k}\widehat{\xi}^N(x_i) - \prod_{i=1}^m \eta^N(x_i)\prod_{i=m+1}^{m+k}\widehat{\eta}^N(x_i)| \le \sum_{i=1}^{m+k}|\xi^N(x_i)-\eta^N(x_i)| \tag{9.10}$$
$$\forall\, \xi^N,\eta^N \in \{0,1\}^{S_N},\ x_i\in S_N,$$

with $m=k=1$, and the coupling $\xi^{N,\mathsf{vm}}_{u_N}\vee\xi^N_{u_N}\le\overline{\xi}^N_{u_N}$, to conclude that for all $s\ge u_N$,

$$\begin{aligned}\Delta_N(s) &\le E_{\xi^N_{s-u_N}}\Big(|G^N(\xi^N_{u_N}) - G^N(\xi^{N,\mathsf{vm}}_{u_N})|\Big)\\ &\le E_{\xi^N_{\xi^N_{s-u_N}}}\Big(\frac{1}{N'}\sum_{e\in\mathcal{N}_N}p_N(e)\sum_{x\in S_N}[(\overline{\xi}^N_{u_N}(x)-\xi^N_{u_N}(x)) + (\overline{\xi}^N_{u_N}(x+e)-\xi^N_{u_N}(x+e))]\\ &\qquad + [(\overline{\xi}^N_{u_N}(x)-\xi^{N,\mathsf{vm}}_{u_N}(x)) + (\overline{\xi}^N_{u_N}(x+e)-\xi^{N,\mathsf{vm}}_{u_N}(x+e))]\Big)\\ &\le 2E_{\xi^N_{s-u_N}}(2\overline{X}^N_{u_N}(\mathbf{1}) - X^N_{u_N}(\mathbf{1}) - X^{N,\mathsf{vm}}_{u_N}(\mathbf{1})) \le C(\log N)^{3-p}X^N_{s-u_N}(\mathbf{1}).\end{aligned} \tag{9.11}$$

The last inequality holds by Lemma 7.1. For small s in (9.7) we will use the crude bound

$$\frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi^N_s(x)\widehat{\xi}^N_s(x+e)\le X^N_s(\mathbf{1}). \tag{9.12}$$

Use (9.8), (9.11) and (9.12) (the latter for $s\le u_N$) in (9.7) to see that

$$\begin{aligned}\mathcal{T}_1 \le& C\delta_N\log N\int_0^{u_N\wedge t}(t-s+\delta_N)^{-1}E(X^N_s(\mathbf{1}))\,ds + C\delta_N\log N\int_{u_N}^{u_N\vee t}(t-s+\delta_N)^{-1}E(\Delta_N(s))\,ds\\ &+ C\delta_N\log N\int_{u_N}^{u_N\vee t}(t-s+\delta_N)^{-1}E\Big(\widehat{E}\Big(\frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi^N_{s-u_N}(B^{N,x}_{u_N})\widehat{\xi}^N_{s-u_n}(B^{N,x+e}_{u_N})\Big)\Big)\,ds\\ \le& C_T\delta_N\log N X^N_0(\mathbf{1})\int_0^{u_N\wedge t}(t-s+\delta_N)^{-1}\,ds + C\delta_N(\log N)^{4-p}X^N_0(\mathbf{1})\int_{u_N}^{u_N\vee t}(t-s+\delta_N)^{-1}\,ds\\ &+ C\delta_N\log N\int_{u_N}^{u_N\vee t}(t-s+\delta_N)^{-1}E\Big(\widehat{E}\Big(\frac{1}{N'}\sum_{e\in\mathcal{N}_N}\sum_{x\in S_N}p_N(e)\xi^N_{s-u_N}(B^{N,x}_{u_N})1(\sigma^N(x,x+e)>u_N)\Big)\Big)\,ds\\ :=& \mathcal{T}_{1,1}+\mathcal{T}_{1,2}+\mathcal{T}_{1,3},\end{aligned} \tag{9.13}$$

where the mean mass bound from Proposition 7.6 (a) is again used in the second inequality. For $s\le u_N(\le\delta_N/2)$ we have $s\le\delta_N/2\le(t+\delta_N)/2$ and so

$$\mathcal{T}_{1,1}\le C_T\delta_N\log N X^N_0(\mathbf{1})u_N 2(t+\delta_N)^{-1}\le C_T X^N_0(\mathbf{1})\frac{\delta_N}{t+\delta_N}. \tag{9.14}$$

We also have

$$\mathcal{T}_{1,2}\le C\delta_N(\log N)^{4-p}X^N_0(\mathbf{1})1(t>u_N)\log\Big(1+\frac{t}{\delta_N}\Big). \tag{9.15}$$

Next use translation invariance of the coalescing walks, Proposition 7.6 (a), and (1.7) to obtain

$$\begin{aligned}\mathcal{T}_{1,3} \le C\delta_N \log N \int_{u_N}^{u_N\vee t} (t-s+\delta_N)^{-1} E\Big(\frac{1}{N'}\sum_{w\in S_N} \xi^N_{s-u_N}(w)\\ \times \sum_{e\in\mathcal{N}_N} p_N(e) \sum_{x\in S_N} \widehat{P}(B^{N,0}_{u_N} = w-x, \sigma^N(0,e) > u_N)\Big)\, ds\\ \le C\delta_N \log N\Big[\sum_{e\in\mathcal{N}_N} p_N(e)\widehat{P}(\sigma^N(0,e)>u_N)\Big]\int_{u_N}^{t\vee u_N}(t-s+\delta_N)^{-1}E(X^N_{s-u_N}(\mathbf{1}))\,ds\\ \le C_T\delta_N X^N_0(\mathbf{1})\frac{\log N}{\log(Nu_N)} 1(t>u_N)\log\Big(\frac{t-u_N+\delta_N}{\delta_N}\Big)\\ \le C_T\delta_N X^N_0(\mathbf{1})\log\Big(1+\frac{t}{\delta_N}\Big). \qquad (9.16)\end{aligned}$$

In the last line we have used $\delta_N \ge cN^{-1/2}$ for N large (which we may assume), and hence the same for u_N. Use (9.14)-(9.16) in (9.13) to conclude

$$\mathcal{T}_1 \le C_T X^N_0(\mathbf{1})\Big[\delta_N\log\Big(1+\frac{t}{\delta_N}\Big)+\frac{\delta_N}{t+\delta_N}\Big]. \qquad (9.17)$$

We next decompose $\mathcal{T}_3^{(j)}$ as

$$\begin{aligned}\mathcal{T}_3^{(j)} \le& \frac{3\delta_N}{N}\sum_{z\in S_N} E\Big(\Big(\int_{u_N}^{t\vee u_N} d^{N,j}(s,\xi^N_s,\phi^z) - E(d^{N,j}(s,\xi^N_s,\phi^z)|\mathcal{F}^N_{s-u_N})\,ds\Big)^2\Big)\\ &+\frac{3\delta_N}{N}\sum_{z\in S_N} E\Big(\Big(\int_{u_N}^{t\vee u_N} E(d^{N,j}(s,\xi^N_s,\phi^z)|\mathcal{F}^N_{s-u_N})\,ds\Big)^2\Big)\\ &+\frac{6\delta_N}{N}\sum_{z\in S_N} E\Big(\int_0^{u_N\wedge t}\Big[\int_0^{s_1} d^{N,j}(s_1,\xi^N_{s_1},\phi^z)d^{N,j}(s_2,\xi^N_{s_2},\phi^z)\,ds_2\Big]\,ds_1\Big)\\ :=&\mathcal{T}^{(j)}_{3,1}+\mathcal{T}^{(j)}_{3,2}+\mathcal{T}^{(j)}_{3,3}. \qquad (9.18)\end{aligned}$$

Equation (7.2) implies,

$$\frac{1}{N'}\sum_{x\in S_N}|d^{N,j}(x,\xi^N_s)| \le \frac{C}{N'}\sum_{x\in S_N}\Big[\Big(\sum_{e\in\bar{\mathcal{N}}_N}\xi^N_s(x+e)\Big)\Big] \le C_{9.19}X^N_s(\mathbf{1}). \qquad (9.19)$$

Recall that

$$d^{N,j}(s,\xi^N_s,\phi^z) = \frac{\ell^{(j)}_N}{N'}\sum_{x\in S_N} p^N_{t-s+\delta_N}(z-x)d^{N,j}(x,\xi^N_s). \qquad (9.20)$$

For $\mathcal{T}^{(j)}_{3,3}$, use (9.20) (take absolute values and use the triangle inequality), do the sum over z first, and then apply Chapman-Kolmogorov and (9.1) to conclude that

$$\begin{aligned}\mathcal{T}^{(j)}_{3,3} \le \frac{6\delta_N}{N}\int_0^{u_N\wedge t}\int_0^{s_1}\frac{(\log N)^6}{(N')^2}NC_{9.1}(2(t+\delta_N)-s_1-s_2)^{-1} \qquad (9.21)\\ \times\sum_{x_1\in S_N}\sum_{x_2\in S_N}E(|d^{N,j}(x_1,\xi^N_{s_1})|\,|d^{N,j}(x_2,\xi^N_{s_2})|)\,ds_2ds_1.\end{aligned}$$

An application of (9.19) and the second moment bound (7.42) now gives

$$\begin{aligned}\mathcal{T}^{(j)}_{3,3} &\le C\delta_N(\log N)^6(X^N_0(\mathbf{1})+X^N_0(\mathbf{1})^2)\int_0^{u_n}\int_0^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}\,ds_2ds_1\\ &\le C\delta_N(\log N)^{6-p}(X^N_0(\mathbf{1})+X^N_0(\mathbf{1})^2), \qquad (9.22)\end{aligned}$$

the last by a bit of calculus.

For $\mathcal{T}_{3,1}^{(j)}$, first use the orthogonality

$$E\Big(\prod_{i=1}^{2}(d^{N,j}(s_i,\xi_{s_i}^N,\phi^z)-E(d^{N,j}(s_i,\xi_{s_i}^N,\phi^z)|\mathcal{F}_{s_i-u_N}^N))\Big)=0 \text{ if } s_2-u_N>s_1$$

to see that

$$\mathcal{T}_{3,1}^{(j)}\le\frac{6\delta_N}{N}\sum_{z\in S_N}E\Big(\int_{u_N}^{u_N\vee t}\Big[\int_{s_1}^{s_1+u_N}\prod_{i=1}^{2}\Big(|d^{N,j}(s_i,\xi_{s_i}^N,\phi^z)|+E\Big(|d^{N,j}(s_i,\xi_{s_i}^N,\phi^z)|\Big|\mathcal{F}_{s_i-u_N}^N\Big)\Big)\,ds_2\Big]ds_1\Big).$$

Argue just as in (9.21), but now with a different product inside the integral, to conclude that

$$\begin{aligned}\mathcal{T}_{3,1}^{(j)}\le&C\delta_N(\log N)^6\int_{u_N}^{u_N\vee t}\Big[\int_{s_1}^{s_1+u_N}\frac{1}{N'}\sum_{x_1\in S_N}\frac{1}{N'}\sum_{x_2\in S_N}(2(t+\delta_N)-s_1-s_2)^{-1}\\&\times E\Big(\prod_{i=1}^{2}\Big[|d^{N,j}(x_i,\xi_{s_i}^N)|+E(|d^{N,j}(x_i,\xi_{s_i}^N)||\mathcal{F}_{s_i-u_N}^N)\Big]\Big)ds_2ds_1.\end{aligned}$$

Now bring the sums through the expectation and product and apply (9.19) to conclude

$$\begin{aligned}\mathcal{T}_{3,1}^{(j)}\le&C\delta_N(\log N)^6\int_{u_N}^{u_N\vee t}\Big[\int_{s_1}^{s_1+u_N}(2(t+\delta_N)-s_1-s_2)^{-1}E\Big(\Big(X_{s_1}^N(\mathbf{1})+E(X_{s_1}^N(\mathbf{1})|\mathcal{F}_{s_1-u_N}^N)\Big)\\&\times\Big(X_{s_2}^N(\mathbf{1})+E(X_{s_2}^N(\mathbf{1})|\mathcal{F}_{s_2-u_N}^N)\Big)\Big)\,ds_2\Big]ds_1\\\le&C\delta_N(\log N)^6\int_{u_N}^{u_N\vee t}\Big[\int_{s_1}^{s_1+u_N}(2(t+\delta_N)-s_1-s_2)^{-1}\\&\times E((X_{s_1}^N(\mathbf{1})+X_{s_1-u_N}^N(\mathbf{1}))(X_{s_2}^N(\mathbf{1})+X_{s_2-u_N}^N(\mathbf{1})))\,ds_2\Big]ds_1,\end{aligned}$$

where in the last we have used the Markov property and mean mass bound from Propostion 7.6. The second moment bound (7.42) from the same result therefore shows that

$$\begin{aligned}\mathcal{T}_{3,1}^{(j)}\le&C_T\delta_N(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)(\log N)^6u_N\int_{u_N}^{u_N\vee t}\frac{1}{2(t-s_1+\delta_N)-u_N}\,ds_1\\\le&C_T\delta_N(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)(\log N)^{6-p}1(t>u_N)\log\Big(\frac{2(t+\delta_N)-3u_N}{2\delta_N-u_N}\Big)\\\le&C_T\delta_N(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)(\log N)^{6-p}\log\Big(\frac{2(t+\delta_N)}{\delta_N}\Big)\quad\text{(recall } u_N\le\delta_N)\\=&C_T(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)(\log N)^{6-p}\delta_N\Big[\log 2+\log\Big(1+\frac{t}{\delta_N}\Big)\Big].\end{aligned}\tag{9.23}$$

Turning to $\mathcal{T}_{3,2}^{(j)}$, for $j=2,3$ introduce

$$H^{N,j}(\xi_0^N,x,u)=E_{\xi_0^N}(d^{N,j}(x,\xi_u^N)),\tag{9.24}$$

and recall from (7.20) that $\hat{H}^{N,j}(\xi_0^N,x,u):=E_{\xi_0^N}(d^{N,j}(x,\xi_u^{N,\mathtt{vm}}))$ satisfies

$$\begin{aligned}\hat{H}^{N,2}(\xi_0^N,x,u_N)&=\sum_{\emptyset\neq A\subset\mathcal{N}_N}r^{N,a}(\sqrt{N}A)\widehat{E}(I^{N,+}(x,u_N,A,\xi_0^N)),\\\hat{H}^{N,3}(\xi_0^N,x,u_N)&=\sum_{\emptyset\neq A\subset\mathcal{N}_N}r^{N,s}(\sqrt{N}A)\widehat{E}((I^{N,+}-I^{N,-})(x,u_N,A,\xi_0^N)).\end{aligned}\tag{9.25}$$

The Markov property of ξ^N and then Chapman-Kolmogorov, imply that for $j=2,3$,

$$\begin{aligned}
&\mathcal{T}_{3,2}^{(j)}\\
&=\frac{6\delta_N}{N}E\Big(\int_{u_N}^{t\vee u_N}\Big[\int_{u_N}^{s_1}\frac{(\ell_N^{(j)})^2}{(N')^2}\sum_{x_1\in S_N}\sum_{x_2\in S_N}\Big[\sum_{z\in S_N}\prod_{i=1}^2 p^N_{t-s_i+\delta_N}(z-x_i)\Big]\prod_{k=1}^2 H^{N,j}(\xi^N_{s_k-u_N},x_k,u_N)\,ds_2\Big]ds_1\Big)\\
&=6\delta_N E\Big(\int_{u_N}^{t\vee u_N}\Big[\int_{u_N}^{s_1}\frac{(\ell_N^{(j)})^2}{(N')^2}\sum_{x_1\in S_N}\sum_{x_2\in S_N}p^N_{2(t+\delta_N)-s_1-s_2}(x_2-x_1)\prod_{k=1}^2 H^{N,j}(\xi^N_{s_k-u_N},x_k,u_N)\,ds_2\Big]ds_1\Big).
\end{aligned}$$

The bound (9.19) and Proposition 7.6(a) show that

$$\frac{1}{N'}\sum_{x\in S_N}|H^{N,j}(\xi_0^N,x,u_N)|\le E_{\xi_0^N}\Big(\frac{1}{N'}\sum_{x\in S_N}|d^{N^j}(x,\xi^N_{u_N})|\Big)\le CE_{\xi_0^N}(X^N_{u_N}(\mathbf{1}))\le C_{9.26}X_0^N(\mathbf{1}). \tag{9.26}$$

Similarly we have

$$\frac{1}{N'}\sum_{x\in S_N}|\hat H^{N,j}(\xi_0^N,x,u_N)|\le C_{9.27}X_0^N(\mathbf{1}). \tag{9.27}$$

Now use (7.16), the coupling from (7.11), and the triangle inequality to conclude that

$$|d^{N,j}(x,\xi^N_{u_N})-d^{N,j}(x,\xi^{N,\mathsf{vm}}_{u_N})|\le C\sum_{y\in x+\bar{\mathcal{N}}_N}(2\overline{\xi}^N_{u_N}(y)-\xi^N_{u_N}(y)-\xi^{N,\mathsf{vm}}_{u_N}(y))\quad\text{for } j=2,3.$$

This implies that

$$\begin{aligned}
\frac{1}{N'}\sum_{x\in S_N}|H^{N,j}(\xi_0^N,x,u_N)-\hat H^{N,j}(\xi_0^N,x,u_N)|&\le CE(2\overline{X}^N_{u_N}(\mathbf{1})-\underline{X}^N_{u_N}(\mathbf{1})-\underline{X}^{N,\mathsf{vm}}_{u_N}(\mathbf{1}))\\
&\le C(\log N)^{3-p}X_0^N(\mathbf{1}).
\end{aligned}\tag{9.28}$$

In the last we used Lemma 7.1. For $j=2,3$, define

$$\hat{\mathcal{T}}_{3,2}^{(j)}=6\delta_N E\Big(\int_{u_N}^{t\vee u_N}\Big[\int_{u_N}^{s_1}\frac{(\ell_N^{(j)})^2}{(N')^2}\sum_{x_1\in S_N}\sum_{x_2\in S_N}p^N_{2(t+\delta_N)-s_1-s_2}(x_2-x_1)\prod_{k=1}^2 \hat H^{N,j}(\xi^N_{s_k-u_N},x_k,u_N)\,ds_2\Big]ds_1\Big).$$

Use (9.1), and then (9.26), (9.27) and (9.28), to conclude

$$\begin{aligned}
|\mathcal{T}_{3,2}^{(j)}-\hat{\mathcal{T}}_{3,2}^{(j)}|&\le C\delta_N\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(\ell_N^{(j)})^2(2(t+\delta_N)-s_1-s_2)^{-1}\\
&\qquad\times\frac{1}{N'}\sum_{x_1\in S_N}\frac{1}{N'}\sum_{x_2\in S_N}\Big|E\Big(\prod_{k=1}^2 H^{N,j}(\xi^N_{s_k-u_N},x_k,u_N)-\prod_{k=1}^2\hat H^{N,j}(\xi^N_{s_k-u_N},x_k,u_N)\Big)\Big|\,ds_2ds_1\\
&\le C\delta_N(\log N)^6\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}E\Big(\Big[\frac{1}{N'}\sum_{x_1\in S_N}|H^{N,j}(\xi^N_{s_1-u_N},x_1,u_N)|\Big]\\
&\qquad\times\Big[\frac{1}{N'}\sum_{x_2\in S_N}|H^{N,j}(\xi^N_{s_2-u_N},x_2,u_N)-\hat H^{N,j}(\xi^N_{s_2-u_N},x_2,u_N)|\Big]\Big)\,ds_2sds_1\\
&\quad+C\delta_N(\log N)^6\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}E\Big(\Big[\frac{1}{N'}\sum_{x_2\in S_N}|\hat H^{N,j}(\xi^N_{s_2-u_N},x_2,u_N)|\Big]\\
&\qquad\times\Big[\frac{1}{N'}\sum_{x_1\in S_N}|H^{N,j}(\xi^N_{s_1-u_N},x_1,u_N)-\hat H^{N,j}(\xi^N_{s_1-u_N},x_1,u_N)|\Big]\Big)\,ds_2sds_1\\
&\le C\delta_N(\log N)^6(\log N)^{3-p}\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}E(X^N_{s_1-u_N}(\mathbf{1})X^N_{s_2-u_N}(\mathbf{1}))ds_2ds_1\\
&\le C_T\delta_N(\log N)^{-2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2)\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}ds_2ds_1.
\end{aligned}$$

In the last line we have again used Proposition 7.6(b). The above integral is uniformly bounded in N and $t \le T$, and so for $j = 2, 3$,

$$|\mathcal{T}_{3,2}^{(j)} - \hat{\mathcal{T}}_{3,2}^{(j)}| \le C_T \delta_N (X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2)(\log N)^{-2}. \tag{9.29}$$

Next consider $\hat{\mathcal{T}}_{3,2}^{(3)}$. For $i \in \{2, ..., |\bar{\mathcal{N}}|\}$ and $\emptyset \ne A \subset \mathcal{N}_N$, recall the notation $I_i^{N,\pm}$ from (7.26), and let

$$\hat{H}_{i,A}^N(\xi_0^N, x, u_N) = \widehat{E}(I_i^{N,+}(x, u_N, A, \xi_0^N) - I_i^{N,-}(x, u_N, A, \xi_0^N)).$$

Then from (9.25) and (7.27) we have

$$\hat{H}^{N,3}(\xi_0^N, x, u_N) = \sum_{i=2}^{|\bar{\mathcal{N}}|} \sum_{\emptyset \ne A \subset \mathcal{N}_N} r^{N,s}(\sqrt{N}A) \hat{H}_{i,A}^N(\xi_0^N, x, u_N). \tag{9.30}$$

Let $i \in \{3, \dots, |\bar{\mathcal{N}}|\}$ and $\emptyset \ne A \subset \mathcal{N}_N$. On $\{B_{u_N}^{N,x+A} \subset \xi_0^N,\ B_{u_N}^{N,x+\bar{\mathcal{N}}_N \setminus A} \subset \widehat{\xi}_0^N,\ |B_{u_N}^{N,x+\bar{\mathcal{N}}_N}| = i\}$, there are distinct points $b_1, b_2, b_3 \in \bar{\mathcal{N}}_N$ so that $\xi_0^N(B_{u_N}^{N,x+b_1}) = 1$ and $\sigma_x^N(b_1, b_2, b_3) > u_N$. The same conclusion holds on $\{B_{u_N}^{N,x+\bar{\mathcal{N}}_N \setminus A} \subset \xi_0^N,\ B_{u_N}^{N,x+A} \subset \widehat{\xi}_0^N,\ |B_{u_N}^{N,x+\bar{\mathcal{N}}_N}| = i\}$. It follows from (3.31) that

$$\begin{aligned} &\sum_{i=3}^{|\bar{\mathcal{N}}|} \sum_{\emptyset \ne A \subset \mathcal{N}_N} |r^{N,s}(\sqrt{N}A) \hat{H}_{i,A}^N(\xi_0^N, x, u_N)| \\ &\qquad \le C \sum_{b_1, b_2, b_3 \in \bar{\mathcal{N}}_N \text{ distinct}} \widehat{E}(\xi_0^N(B_{u_N}^{N,x+b_1}) 1(\sigma_x^N(b_1, b_2, b_3) > u_N)). \end{aligned} \tag{9.31}$$

In the right-hand side of the above, sum over the possible values of $B_{u_N}^{N,x+b_1}$, as in the proof of Lemma 7.8(a), and then use (1.7) with $n = 3$ and (9.1) to conclude that

$$\begin{aligned} &\frac{1}{N'} \sum_{x_2 \in S_N} p_{2(t+\delta_N)-s_1-s_2}^N(x_2 - x_1) \Big[\sum_{i=3}^{|\bar{\mathcal{N}}|} \sum_{\emptyset \ne A \subset \mathcal{N}_N} |r^{N,s}(\sqrt{N}A) \hat{H}_{i,A}^N(\xi_{s_2-u_N}^N, x_2, u_N)| \Big] \\ &\quad \le C_{9.32} (\log N)^{-3} (2(t+\delta_N) - s_1 - s_2)^{-1} X_{s_2-u_N}^N(\mathbf{1}). \end{aligned} \tag{9.32}$$

The same reasoning as above gives for all $s \ge u_N$,

$$\frac{1}{N'} \sum_{x \in S_N} \Big[\sum_{i=3}^{|\bar{\mathcal{N}}|} \sum_{\emptyset \ne A \subset \mathcal{N}_N} |r^{N,s}(\sqrt{N}A) \hat{H}_{i,A}^N(\xi_{s-u_N}^N, x, u_N)| \Big] \le C_{9.33} (\log N)^{-3} X_{s-u_N}^N(\mathbf{1}). \tag{9.33}$$

Next consider the contribution to $\hat{\mathcal{T}}_{3,2}^{(3)}$ from the $i = 2$ term in (9.30). For $u_N \le s_i \le t$, $x_i \in S_N$ ($i = 1, 2$), and $k = 1, 2$, define

$$\Phi^{(s_{3-k}, x_{3-k})}(s_k, x_k) = p_{2(t+\delta_N)-s_k-s_{3-k}}^N(x_k - x_{3-k}) = p_{2(t+\delta_N)-s_1-s_2}^N(x_2 - x_1).$$

By Lemma 2.1 of [9], we have

$$\|\Phi^{s_{3-k} x_{3-k}}(s_k, \cdot)\|_{\text{Lip}} \le C_{9.34} (2(t+\delta_N) - s_1 - s_2)^{-3/2}. \tag{9.34}$$

For $k = 1, 2$, by the definitions of $\hat{H}_2^{N,3}$ in (7.29) and $\hat{H}_{2,A}^N$, and (7.28) we have,

$$\begin{aligned} &\frac{1}{N'} \Big| \sum_{x_k \in S_N} p_{2(t+\delta_N)-s_1-s_2}^N(x_2 - x_1) \sum_{\emptyset \ne A \subset \mathcal{N}} r^{N,s}(\sqrt{N}A) \hat{H}_{2,A}^N(\xi_{s_k-u_N}^N, x_k, u_N) \Big| \\ &\quad = (\log N)^{-3} |\hat{H}_2^{N,3}(\xi_{s_k-u_N}^N, u_N, \Phi_{s_k}^{(s_{3-k}, x_{3-k})})| \\ &\quad \le C_{9.35} (2(t+\delta_N) - s_1 - s_2)^{-3/2} X_{s_k-u_N}^N(\mathbf{1}) (\log N)^{-p/2}, \end{aligned} \tag{9.35}$$

where in the last line we used (7.33) and (9.34). For each $\emptyset \neq A \subset \mathcal{N}_N$ we choose $a \in A$. A much cruder calculation than that above shows that,

$$\begin{aligned}
&\frac{1}{N'}\sum_{x_1\in S_N}\sum_{\emptyset\neq A\subset\mathcal{N}}|r^{N,s}(\sqrt{N}A)\hat{H}^N_{2,A}(\xi^N_{s_1-u_N},x_1,u_N)| \\
&\quad\le \frac{\|r\|}{N'}\sum_{x_1\in S_N}\sum_{\emptyset\neq A\subset\mathcal{N}}\widehat{E}\Big(\Big(\xi^N_{s_1-u_N}(B^{N,x_1+a}_{u_N})+\xi^N_{s_1-u_N}(B^{N,x_1}_{u_N})\Big) \\
&\qquad\qquad\times 1(|B^{N,x_1+A}_{u_N}|=|B^{N,x_1+\bar{\mathcal{N}}_N\setminus A}_{u_N}|=1, B^{N,x_1+a}_{u_N}\neq B^{N,x_1}_{u_N})\Big) \\
&\quad\le \frac{\|r\|}{N'}\sum_{w\in S_N}\xi^N_{s_1-u_N}(w)\sum_{x_1\in S_N}\Big[\widehat{P}(B^{N,a}_{u_N}=w-x_1,\sigma^N(0,a)>u_N) \\
&\qquad\qquad+\widehat{P}(B^{N,0}_{u_N}=w-x_1,\sigma^N(0,a)>u_N)\Big] \\
&\quad= 2\|r\|X^N_{s_1-u_N}(\mathbf{1})\widehat{P}(\sigma^N(0,a)>u_N)\le C_{9.36}X^N_{s_1-u_N}(\mathbf{1})(\log N)^{-1}, \qquad (9.36)
\end{aligned}$$

the last by (1.7) with $n=2$ and $u_N \ge cN^{-1/2}$ (by(9.2)). Now argue exactly as in the derivation of (89) in [9] following the derivation there from just below (88) to (89). Here the hypotheses (85)-(88) of that argument are provided by (9.35), (9.36), (9.32) and (9.33), respectively. From this derivation we may conclude that

$$\begin{aligned}
|\hat{\mathcal{T}}^{(3)}_{3,2}| \le & 6\delta_N(\log N)^6\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}\Big|E\Big(\frac{1}{N'}\sum_{x_1\in S_N}\frac{1}{N'}\sum_{x_2\in S_N}p^N_{2(t+\delta_N)-s_1-s_2}(x_2-x_1) \\
&\times\prod_{k=1}^2\Big[\sum_{i=2}^{|\bar{\mathcal{N}}|}\sum_{\emptyset\neq A\subset\mathcal{N}_N}r^{\mathbb{N},s}(\sqrt{N}A)\hat{H}^N_{i,A}(\xi^N_{s_k-u_N},x_k,u_N)\Big]\Big)\Big|\,ds_2ds_1 \\
\le & C\delta_N\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}\frac{E(X^N_{s_1-u_N}(\mathbf{1})X^N_{s_2-u_N}(\mathbf{1}))}{(2(t+\delta_N)-s_1-s_2)^{3/2}}ds_2ds_1(\log N)^{6-p/2}((\log N)^{-1}+(\log N)^{-3}) \\
&+C\delta_N\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}E(X^N_{s_1-u_N}(\mathbf{1})X^N_{s_2-u_N}(\mathbf{1}))(2(t+\delta_N)-s_1-s_2)^{-1}ds_2ds_1 \\
\le & C_T\delta_N(X^N_0(\mathbf{1})+X^N_0(\mathbf{1})^2). \qquad (9.37)
\end{aligned}$$

In the last line we have used (7.42).

The corresponding bound on $\hat{\mathcal{T}}^{(2)}_{3,2}$ is much simpler. The analogue of (9.31) is

$$\sum_{i=2}^{\bar{\mathcal{N}}}\sum_{\emptyset\neq A\subset\mathcal{N}}|r^{N,a}(\sqrt{N}A)|\hat{E}(I^{N,+}_i(x,u_N,A,\xi^N_0))\le CX^N_0(1)(\log N)^{-1}.$$

Now proceed as in (9.32) using the supnorm bound on $p^N_{2(t+\delta_N-s_1-s_2}(x_1-x_2)$ to see that

$$\begin{aligned}
|\hat{\mathcal{T}}^{(2)}_{3,2}| &\le C\delta_N\int_{u_N}^{t\vee u_N}\int_{u_N}^{s_1}(2(t+\delta_N)-s_1-s_2)^{-1}E(X^N_{s_1-u_N}(1)X^N_{s_2-u_N}(1))\,ds_1ds_2 \\
&\le C_T\delta_N(X^N_0(1)+X^N_0(1)^2). \qquad (9.38)
\end{aligned}$$

Finally use the decompositions (9.3) and (9.18) along with the bounds on $\mathcal{T}_0$, $\mathcal{T}_2$, $\mathcal{T}_1$, $\mathcal{T}_{3.3}$, $\mathcal{T}_{3.1}$, $|\mathcal{T}^{j)}_{3.2}-\hat{\mathcal{T}}^{(j)}_{3.2}|$, and $|\hat{\mathcal{T}}^{(j)}_{3.2}|$ from (9.4), (9.6), (9.17), (9.22), (9.23), (9.29), (9.37) and (9.38), respectively, to see that

$$E\Big(\int\int 1_{\{|x-y|\le\sqrt{\delta_N}\}}dX^N_t(x)dX^N_t(y)\Big)\le C_T(X^N_0(\mathbf{1})+X^N_0(\mathbf{1})^2)\delta_N\Big(1+\log\Big(1+\frac{t}{\delta_N}\Big)+(t+\delta_N)^{-1}\Big).$$

The result follows by noting that $\log\Big(1+\frac{t}{\delta_N}\Big)+(t+\delta_N)^{-1}$ is bounded away from 0 uniformly in N, $t>0$ and so we can drop the initial 1 on the right-hand side. $\square$

10 Proof of Theorem 4.9

We assume the hypotheses of Theorem 4.9, and work in the setting of Sections 4 and 6, so that the monotone, asymptotically symmetric voter model perturbation, $\{\xi^{[\varepsilon]}: 0<\varepsilon\le\varepsilon_0\}$ is constructed as in Proposition (4.3)(a), along with it's associated measure-valued process X^N in (6.5). For real numbers $K_0>2$ and $L'>3$, let $I=[-L',L']^2$, $\tilde I=(-K_0L',K_0L')^2$, and recall the notation $I_{\pm e_i}=\pm L'e_i+[-L'+1,L'-1]^2$, $i=1,2$, from (4.18). Recall also that N is chosen as in (3.16). We construct our killed processes $\underline{\xi}^{[\varepsilon]}$ as in (4.5), where killing is done when $|x|\ge M_0:=\lfloor\sqrt{N}K_0L'\rfloor$ ($x\in\mathbb{Z}^2$). We define $\underline{\xi}^N_t(x)=\underline{\xi}^{[\varepsilon_N]}_{Nt}(\sqrt{N}x)$ for $x\in S_N$ as in (4.16), and let $\xi^N_t(x)=\xi^{[\varepsilon_N]}_{Nt}(\sqrt{N}x)$ for $x\in S_N$ as in Section 6. The killed measure valued process $\underline{X}^N$ is defined as in (4.17) and so the killing here is done for $|x|\ge K_0L'$ ($x\in S_N$). By Proposition 4.3 (a),(b) our processes are therefore coupled so that

$$\underline{\xi}^N\le\xi^N \text{ and hence } \underline{X}^N\le X^N. \tag{10.1}$$

Recall from Section 7 that $(P^N_t,t\ge0)$ is the semigroup of a rate N random walk on S_N with step kernel p_N. For $x\in S_N$ let $\tilde B^{N,x}$ denote a random walk starting at $x\in S_N$ with semigroup (P^N_t).

A key step in verifying condition (4.18) is to show that the killed and unkilled processes are close on certain time scales through the following version of Lemma 8.1 of [9]. We stress that for now K_0 and L' are arbitrary real parameters.

Lemma 10.1. *There is a positive constant $c_{10.1}$ and a nondecreasing function $C_{10.1}:\mathbb{R}_+\to\mathbb{R}_+$ so that if $t>0$, $K_0>2$ and $L'>3$, and $\underline{X}^N_0=X^N_0$ is supported on I, then*

$$\begin{aligned}E[X^N_t(\mathbf{1})-\underline{X}^N_t(\mathbf{1})]\le X^N_0(\mathbf{1})\Big[&c_{10.1}e^{c_{10.1}t}P\Big(\sup_{s\le t}|\tilde B^{N,0}_s|>(K_0-1)L'-3\Big)\\&+C_{10.1}(t)(1\vee X^N_0(\mathbf{1}))(\log N)^{-1/6}\Big].\end{aligned}\tag{10.2}$$

The Lemma is proved below but we first turn to the main result of this Section. Given Lemma 10.1, the proof of Theorem 4.9 is done just as the proof of Lemma 6.2 of [13] (for Lotka-Volterra models). We outline the argument below for completeness.

Proof of Theorem 4.9 (sketch). Recall that we must show, after perhaps reducing ε_0, that

$$\begin{aligned}&\text{There are } T'>1,\ K,J'\in\mathbb{N} \text{ with } K>2, \text{ and } L'>3, \text{ so that if } 0<\varepsilon\le\varepsilon_0,\\&\text{then for } L=\lfloor\sqrt{N}L'\rfloor,\ \underline{X}^N_0([-L',L']^2)\ge J' \text{ implies}\\&P(\underline{X}^N_{T'}(I_e)\ge J' \text{ for all } e\in\{\pm e_i,i=1,2\})\ge 1-6^{-5(2K+1)^3}.\end{aligned}\tag{10.3}$$

Note first that our hypotheses imply that Theorem 6.1 holds. The limiting super-Brownian motion in that result has drift $\Theta_2+\Theta_3$, which is positive by hypothesis, and therefore will continue to grow exponentially up to time T' with high probability if it has a large enough initial mass. It is therefore not hard to prove an analogue of (10.3) for this limiting process (see (6.7) of [13]). By the convergence theorem (Theorem 6.1), the same bound will hold for $X^N_{T'}$ if N is large enough, and so ε is sufficiently small. This is where we may need to reduce ε_0. Here we also use monotonicity to reduce to the case where X^N_0 is finite and apply a subsequence argument to assume these initial measures converge and so the convergence theorem holds. To derive the same bound for $\underline{X}^N$, and hence gain the necessary spatial independence required for our comparison to oriented percolation, we need to show $\underline{X}^N_{T'}$ is close to $X^N_{T'}$. This is where Lemma 10.1 is needed. The inputs required to carry out the proof of Lemma 6.2 of [13] are Lemma 10.1 and the weak convergence of the rescaled process to SBM with a positive drift, given here by Theorem 6.1. Our definition of I_e for $e = \pm e_i$ is slightly different from that in [13] but it results in only a trivial change. Also we have been a bit more careful in choosing integer parameters here. So once $K \in \mathbb{N}$ and $L' > 3$ are chosen as in the proof of Lemma 6.2 of [13], one takes N large and sets $L = \lfloor \sqrt{N} L' \rfloor$ (as in (10.3)), and then chooses $K_0 = K_0(N)(> 2)$ so that $K_0 L' \sqrt{N} = KL$. (In this way K_0 is comparable to K.) This equality ensures killing for the unscaled process outside $(-KL, KL)^2$ corresponds to killing $\underline{X}^N$ outside $\tilde{I}$ as in Lemma 10.1. (Note that K_0 may depend on N in Lemma 10.1.) The rest of the argument is now identical to that of Lemma 6.2 in [13]. $\square$

Recall from Section 7 that $\{B^{N,x} : x \in S_N\}$ is a system of rate $w_N N$ coalescing random walks in S_N with step kernel p_N. Let $T'_x = \inf\{t \geq 0 : B^{N,x}_t \notin \tilde{I}\}$, let Δ denote a cemetery state, and define a "killed" coalescing random walk system, $\{\underline{B}^{N,x}_t, x \in S_N\}$, by

$$\underline{B}^{N,x}_t = \begin{cases} B^{N,x}_t & \text{if } t < T'_x \\ \Delta & \text{if } t \geq T'_x. \end{cases}$$

We define the killed random walk $\underline{\tilde{B}}^{N,x}$ (recall the step rate here is N and there is no coalescing) in the same way and denote its associated killed semigroup by $(\underline{P}^N_t, t \geq 0)$. Of course, $\underline{B}^{N,x}_t = \underline{\tilde{B}}^{N,x} \equiv \Delta$ for all $x \notin \tilde{I}$. We will use the convention that $\xi(\Delta) = 0$ for all $\xi \in \{0,1\}^{S_N}$.

Proof of Lemma 10.1. We follow the proof of Lemma 8.1 in [9] for 2-dimensional Lotka-Volterra models, but some modifications are needed. Assume X^N_0 (and hence $\xi^N_0 = \underline{\xi}^N_0$) is supported on $I = [-L', L']^2$ and $T' > 0$. Let $f : S_N \cup \{\Delta\} \to \mathbb{R}$ with $f(\Delta) = 0$ and set $\Phi(s,x) = \underline{P}^N_{t-s} f(x)$, $s \leq t$. We will assume $t \in [0, T']$ in what follows. The killed analogue of (6.10) is derived as in the proof of Lemma 3.2 of [11] where a general class of voter model perturbations is considered. The argument there uses a different representation for the perturbation but applies to our representations without change and gives (see the last display on p. 113 of [11])

$$\underline{X}^N_t(f) = \underline{X}^N_0(\underline{P}^N_t f) + \int_0^t \sum_{j=2}^3 d^{N,j}(s, \underline{\xi}^N_s, \Phi) ds + \underline{M}^N_t(\Phi), \tag{10.4}$$

where $\underline{M}^N_t(\Phi)$ is a square-integrable, mean zero martingale. Next, choose $h : \mathbb{R}^2 \cup \{\Delta\} \to [0,1]$ such that $h(\Delta) = 0$ and

$$[-K_0L' + 3, K_0L' - 3]^2 \subset \{h = 1\} \subset \mathrm{Supp}(h) \subset [-K_0L' + 2, K_0L' - 2]^2, \quad |h|_{\mathtt{Lip}} \leq 1,$$

and define, for $s \leq t$ and $x \in S_N$, $\Psi(s,x) = \underline{P}^N_{t-s} h(x)$. By Lemma 8.4 in [9] there is a constant $C_{10.5} > 0$ such that

$$\|\Psi\|_N \leq C_{10.5} \quad \text{for all } N. \tag{10.5}$$

By (7.103) (with $c = 0$ and $\Phi = \mathbf{1}$), (10.4) (with $\Phi = \Psi$ and $f = h$), the inequality $h \leq 1$ and $X^N_0 = \underline{X}^N_0$, we have

$$\begin{aligned} E[X^N_t(\mathbf{1}) - \underline{X}^N_t(\mathbf{1})] &\leq E[X^N_t(\mathbf{1}) - \underline{X}^N_t(h)] \\ &= X^N_0(\mathbf{1} - \underline{P}^N_t(h)) + E\Big[\int_0^t \sum_{j=2}^3 (d^{N,j}(s, \xi^N_s, \mathbf{1}) - d^{N,j}(s, \underline{\xi}^N_s, \Psi))\, ds\Big]. \end{aligned} \tag{10.6}$$

It follows that

$$E[X_t^N(\mathbf{1}) - \underline{X}_t^N(\mathbf{1})] \le \mathcal{U}_0 + \mathcal{U}_1 + \mathcal{U}_2 + \mathcal{U}_3 + \mathcal{U}_4, \tag{10.7}$$

where

$$\begin{aligned}
\mathcal{U}_0 &= X_0^N(\mathbf{1} - \underline{P}_t^N(h)),\\
\mathcal{U}_1 &= E\Big[\int_0^{t_N\wedge t} \sum_{j=2}^3 (d^{N,j}(s,\xi_s^N,\mathbf{1}) - d^{N,j}(s,\underline{\xi}_s^N,\Psi))\,ds\Big],\\
\mathcal{U}_2 &= E\Big[\int_{t_N}^{t\vee t_N} \sum_{j=2}^3 d^{N,j}(s,\xi_s^N,\mathbf{1}-\Psi)\,ds\Big],\\
\mathcal{U}_3 &= E\Big[\int_{t_N}^{t\vee t_N} (d^{N,2}(s,\xi_s^N,\Psi) - d^{N,2}(s,\underline{\xi}_s^N,\Psi))\,ds\Big],\\
\mathcal{U}_4 &= E\Big[\int_{t_N}^{t\vee t_N} (d^{N,3}(s,\xi_s^N,\Psi) - d^{N,3}(s,\underline{\xi}_s^N,\Psi))\,ds\Big].
\end{aligned}$$

The labeling matches that of (127) in [9].

We claim that there is a positive function $C_{10.9} : (0,\infty) \to (0,\infty)$ and positive constants $\tilde{K}$, $c_{10.10}$ such that for any $t \le T'$, if $|\cdot|_\infty$ denotes the L^∞ norm on $\mathbb{R}^d$, then

$$|\mathcal{U}_0| \le X_0^N(\mathbf{1}) P\Big(\sup_{u\le t} |\tilde{B}_u^{N,0}|_\infty > (K_0-1)L' - 3\Big), \tag{10.8}$$

$$|\mathcal{U}_1| \le C_{10.9}(T') X_0^N(\mathbf{1})(\log N)^3 t_N, \tag{10.9}$$

$$\begin{aligned}
|\mathcal{U}_2| \le X_0^N(\mathbf{1})\Big[&c_{10.10} e^{c_{10.10}t} \hat{P}\Big(\sup_{u\le t} |\tilde{B}_u^{N,0}|_\infty > (K_0-1)L' - 3\Big)\\
&+ C_{10.9}(T')(1 \vee X_0^N(\mathbf{1}))(\log N)^{-1/2}\Big],
\end{aligned} \tag{10.10}$$

$$|\mathcal{U}_j| \le C_{10.9}(T')(\log N)^{-1/6}(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2) + \tilde{K}\int_0^t E[X_s^N(\mathbf{1}) - \underline{X}_s^N(\mathbf{1})]ds,\ j = 3,4. \tag{10.11}$$

Assuming (10.8)–(10.11) for now, and recalling that $t_N = (\log N)^{-19}$, we see that for some function $C : (0,\infty) \to (0,\infty)$, and all $t \le T'$,

$$\begin{aligned}
E[X_t^N(\mathbf{1}) - \underline{X}_t^N(\mathbf{1})] \le X_0^N(\mathbf{1})\Big[&c_{10.10} e^{c_{10.10}t} P\Big(\sup_{s\le t} |\tilde{B}_s^{N,0}|_\infty > (K_0-1)L' - 3\Big)\\
&+ C(T')(1 \vee X_0^N(\mathbf{1}))(\log N)^{-1/6}\Big] + 2\tilde{K}\int_0^t E[X_s^N(\mathbf{1}) - \underline{X}_s^N(\mathbf{1})]ds.
\end{aligned}$$

By replacing $C(T')$ with $\inf_{T\ge T'} C(T)$ we may assume that $C(\cdot)$ is non-decreasing. Now take $T' = t$ in the above and use Gronwall's inequality to complete the proof of Lemma 10.1.

Thus, our remaining task in this section is to verify (10.8)–(10.11). Equation (10.8) is (128) of [9] (which uses the fact that X_0^N is supported on I and so applies here as well). Equation (10.9) follows from (8.2) and its counterpart for $\underline{\xi}^N$ (the proof is the same), the mean mass bound Proposition 7.6(a), and $\underline{X}^N \le X^N$.

The derivation of (10.10) follows that of (130) in [9]. Since it does not involve $\underline{\xi}_t^N$ we can apply the bounds from Section 7. We have

$$\begin{aligned}|\mathcal{U}_2| \leq & \Big|E\Big(\int_{t_N}^{t\vee t_N}\sum_{j=2}^{3}\big(E(d^{N,j}(s,\xi_s^N,1-\Psi)|\mathcal{F}_{s-t_N})-\Theta_j^N X_{s-t_N}^N(1-\Psi_{s-t_N})\big)\,ds\Big)\Big| \\ & +\Big|E\Big(\int_{t_N}^{t\vee t_N}(\Theta_2^N+\Theta_3^N)X_{s-t_N}^N(1-\Psi_{s-t_N})\,ds\Big)\Big| \\ :=& |\Delta_1|+|\Delta_2|. \end{aligned} \tag{10.12}$$

For Δ_1, first use Proposition 7.12 and (10.5), and then Corollary 7.15 and Proposition 7.6(a) to see that for some $C_{10.13}(T')$ (and all $t \leq T'$)

$$|\Delta_1| \leq C_{10.13}(T')(\log N)^{-1/2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2). \tag{10.13}$$

For Δ_2, use (7.66), (7.76), and Lemma 7.19 (recall $\|\Psi_{s-t_N}\|_{\mathtt{Lip}} \leq C_{10.5}$ by (10.5) and take $T' \geq C_{10.5}$ to apply Lemma 7.19), and then use $P_{s-t_N}^N\Psi_{s-t_N} \geq \underline{P}_{s-t_N}^N\Psi_{s-t_N} = \underline{P}_t^N h$ to see that for some constants $c_i > 0$,

$$\begin{aligned}|\Delta_2| &\leq c_1\int_{t_N}^{t\vee t_N} e^{c_{7.101}(t-s)}X_0^N(P_{s-t_N}^N(1-\Psi_{s-t_N}))\,ds + C_{7.101}(T')(\log N)^{-1/2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2) \\ &\leq c_1\int_{t_N}^{t\vee t_N} e^{c_{7.101}(t-s)}X_0^N(1-\underline{P}_t^N h)\,ds + C_{7.101}(T')(\log N)^{-1/2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2) \\ &\leq c_2 e^{c_{7.101}t}X_0^N(\mathbf{1})P(\sup_{u\leq t}|\tilde{B}_u^{N,0}|_\infty > (K_0-1)L'-3) + C_{7.101}(T')(\log N)^{-1/2}(X_0^N(\mathbf{1})+X_0^N(\mathbf{1})^2), \end{aligned} \tag{10.14}$$

where (10.8) is used in the last line. So by (10.12)-(10.14), we have (10.10).

We turn now to the more involved proof of (10.11). Recall from Section 7.1, that $\xi_t^{N,\mathtt{vm}}$ denotes a rescaled voter model on S_N with rate function $Nw_Nc^{N,\mathtt{vm}}(x,\xi)$. Let $\underline{\xi}^{N,\mathtt{vm}}$ denote the corresponding killed voter model, which has rate function $Nw_Nc^{N,\mathtt{vm}}(x,\xi)1(x\in\tilde{I})$ and initial condition $\underline{\xi}_0^N$ supported on $\tilde{I}$. We will assume $\xi^{N,\mathtt{vm}}$ has the same initial condition and so by the monotonicity of the voter model, just as in (10.1), we may assume

$$\underline{\xi}^{N,\mathtt{vm}} \leq \xi^{N,\mathtt{vm}}. \tag{10.15}$$

We will also use the following killed duality equation which is a special case of (9.36) in [12]:

For $\underline{\xi}_0^N$ supported in $\tilde{I}$ and finite disjoint $A,B\subset S_N$,

$$E_{\underline{\xi}_0^N}\Big[\prod_{a\in A}\underline{\xi}_t^{N,\mathtt{vm}}(a)\prod_{b\in B}(1-\underline{\xi}_t^{N,\mathtt{vm}}(b))\Big] = \widehat{E}\Big[\prod_{a\in A}\underline{\xi}_0^N(\underline{B}_t^{N,a})\prod_{b\in B}(1-\underline{\xi}_0^N(\underline{B}_t^{N,b}))\Big]. \tag{10.16}$$

In view of the above we assume for now that $\underline{\xi}_0^N$ is supported on the larger set $\tilde{I}$. Recalling (9.25), for $u>0$ and $j=2,3$ we define

$$\underline{\hat{H}}^{N,j}(\underline{\xi}_0^N,x,u) = E_{\underline{\xi}_0^N}(d^{N,j}(x,\underline{\xi}_u^{N,\mathtt{vm}})). \tag{10.17}$$

By (10.16) and just as in the derivation of (7.20),

$$\begin{aligned}\underline{\hat{H}}^{N,2}(\underline{\xi}_0^N,x,u) &= \sum_{\emptyset\neq A\subset\mathcal{N}_N} r^{N,a}(\sqrt{N}A)\widehat{E}\Big(\prod_{y\in A}\underline{\xi}_0^N(\underline{B}_u^{N,x+y})\prod_{z\in\mathcal{N}_N\setminus A}(1-\underline{\xi}_0^N(\underline{B}_u^{N,x+z}))\Big) \\ \underline{\hat{H}}^{N,3}(\underline{\xi}_0^N,x,u) &= \sum_{\emptyset\neq A\subset\mathcal{N}_N} r^{N,s}(\sqrt{N}A)\widehat{E}\Big(\prod_{y\in A}\underline{\xi}_0^N(\underline{B}_u^{N,x+y})\prod_{z\in\mathcal{N}_N\setminus A}(1-\underline{\xi}_0^N(\underline{B}_u^{N,x+z})) \\ &\qquad -\prod_{y\in A}(1-\underline{\xi}_0^N(\underline{B}_u^{N,x+y}))\prod_{z\in\mathcal{N}_N\setminus A}\underline{\xi}_0^N(\underline{B}_u^{N,x+z})\Big).\end{aligned}$$

With the definitions (7.21) and (7.22) in mind we also introduce

$$\underline{\hat{H}}^{N,j}(\underline{\xi}_0^N, u, \Psi_s) = \frac{\ell_N^{(j)}}{N'} \sum_{x \in S_N} \Psi(s,x) \underline{\hat{H}}^{N,j}(\underline{\xi}_0^N, x, u),\ j = 2, 3.$$

As for (7.23), (10.16) implies for $j = 2, 3$,

$$E_{\underline{\xi}_0^N}[d^{N,j}(s, \underline{\xi}_u^{N,\mathsf{vm}}, \Psi)] = \underline{\hat{H}}^{N,j}(\underline{\xi}_0^N, u, \Psi_s).$$

The next result is a killed version of Lemma 7.3. We give the proof at the end of this section.

Lemma 10.2. *There is a constant $C_{10.18}$ such that for $j = 2, 3$, all $T' > 0$, $\Phi \in C_b([0,T'] \times \mathbb{R}^2)$ and all $s \in [t_N, T']$,*

$$\Big| E_{\underline{\xi}_0^N}\big[d^{N,j}(s, \underline{\xi}_s^N, \Phi) | \mathcal{F}_{s-t_N}^N\big] - \underline{\hat{H}}^{N,j}(\underline{\xi}_{s-t_N}^N, t_N, \Phi_{s-t_N}) \Big| \le C_{10.18} \|\Phi\|_{1/2,N} (\log N)^{-13/2} \underline{X}_{s-t_N}^N(\mathbf{1}). \tag{10.18}$$

Let us assume again that $\xi_0^N = \underline{\xi}_0^N$ is supported on the smaller set I (as opposed to $\tilde{I}$), and for $j = 2, 3$ write $\mathcal{U}_{j+1} = \sum_{i=1}^5 \mathcal{V}_{j,i}$, where

$$\mathcal{V}_{j,1} = E\Big[\int_{t_N}^{t_N \vee t} [d^{N,j}(s, \xi_s^N, \Psi) - \Theta_j^N X_{s-t_N}^N(\Psi)] ds\Big], \tag{10.19}$$

$$\mathcal{V}_{j,2} = \Theta_j^N E\Big[\int_{t_N}^{t_N \vee t} [X_{s-t_N}(\Psi) - \underline{X}_{s-t_N}(\Psi)] ds\Big], \tag{10.20}$$

$$\mathcal{V}_{j,3} = E\Big[\int_{t_N}^{t_N \vee t} [\Theta_j^N \underline{X}_{s-t_N}(\Psi) - \hat{H}^{N,j}(\underline{\xi}_{s-t_N}^N, t_N, \Psi_{s-t_N})] ds\Big], \tag{10.21}$$

$$\mathcal{V}_{j,4} = E\Big[\int_{t_N}^{t_N \vee t} [\hat{H}^{N,j}(\underline{\xi}_{s-t_N}^N, t_N, \Psi_{s-t_N}) - \underline{\hat{H}}^{N,j}(\underline{\xi}_{s-t_N}^N, t_N, \Psi_{s-t_N})] ds\Big], \tag{10.22}$$

$$\mathcal{V}_{j,5} = E\Big[\int_{t_N}^{t_N \vee t} [\underline{\hat{H}}^{N,j}(\underline{\xi}_{s-t_N}^N, t_N, \Psi_{s-t_N}) - d^{N,j}(s, \underline{\xi}_s^N, \Psi)] ds. \tag{10.23}$$

We bound these one at a time.

By (7.91) and (7.98),

$$\begin{aligned}
|\mathcal{V}_{j,1}| &\le C_{7.91} \|\Psi\|_N \int_0^{(t-t_N)^+} \Big(\frac{1}{t_N \log N} E[\mathscr{I}^N(\xi_s^N)] + (\log N)^{-1/2} E[X_s^N(\mathbf{1})] \Big) ds & (10.24)\\
&\le C_{7.91} \|\Psi\|_N \Big[C_{7.98}(T') (\log N)^{-1/2} \Big(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2 \Big) + (\log N)^{-1/2} \int_0^{(t-t_N)^+} E[X_s^N(\mathbf{1})] ds \Big] \\
&\le C_{10.25}(T') (\log N)^{-1/2} \Big(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2 \Big), & (10.25)
\end{aligned}$$

where we have used Proposition 7.6(a) and (10.5) in the last step. Recalling that $\underline{\xi}_t^N \le \xi_t^N$, letting $\tilde{K} = \sup_N (|\Theta_2^N| + |\Theta_3^N|) < \infty$ and using $\|\Psi\|_\infty \le 1$, we get

$$|\mathcal{V}_{j,2}| \le \tilde{K} \int_0^{(t-t_N)^+} E[X_s(\mathbf{1}) - \underline{X}_s(\mathbf{1})] ds. \tag{10.26}$$

Change variables in $\mathcal{V}_{j,3}$ to rewrite it as an integral over $[0, (t-t_N)^+]$, set $u = s$ and replace ξ_0^N with $\underline{\xi}_s^N$ in Proposition 7.10 and Proposition 7.11 to obtain

$$|\mathcal{V}_{j,3}| \le C(T') \|\Psi\|_N \int_0^{(t-t_N)^+} \Big(\frac{1}{t_N \log N} E[\mathscr{I}^N(\underline{\xi}_s^N)] + (\log N)^{-1/2} E[\underline{X}_s^N(\mathbf{1})] \Big) ds,$$

which is the same as the right-hand side of (10.24), but with $\underline{\xi}_s^N$ instead of ξ_s^N. Since $\underline{\xi}_s^N \le \xi_s^N$, (10.25) gives

$$|\mathcal{V}_{j,3}| \le C_{10.25}(\log N)^{-1/2}\Big(X_0^N(\mathbf{1}) + X_0^N(\mathbf{1})^2\Big). \tag{10.27}$$

By (10.5), Lemma 10.2, $\underline{X}^N(\mathbf{1}) \le X^N(\mathbf{1})$ and Proposition 7.6(a),

$$|\mathcal{V}_{j,5}| \le C_{10.18}\|\Psi\|_{1/2,N}(\log N)^{-13/2}\int_0^{(t-t_N)^+} E[\underline{X}_s^N(\mathbf{1})]ds \le C_{10.28}(T')(\log N)^{-13/2}X_0^N(\mathbf{1}). \tag{10.28}$$

Turning to $\mathcal{V}_{j,4}$, we use $\Psi_s \equiv 0$ on $(\tilde{I})^c$ and $\|\Psi\|_\infty \le 1$ to obtain

$$\Big|\hat{H}^{N,j}(\underline{\xi}_s^N, t_N, \Psi_s) - \underline{\hat{H}}^{N,j}(\underline{\xi}_s^N, t_N, \Psi_s)\Big| \le \frac{(\log N)^3}{N'}\sum_{x\in\tilde{I}}|\hat{H}^{N,j}(\underline{\xi}_s^N, x, t_N) - \underline{\hat{H}}^{N,j}(\underline{\xi}_s^N, x, t_N)|. \tag{10.29}$$

For $x \in S_N$ we may use (7.16), the coupling (10.15), and then duality (recall (10.16)) to see that

$$\begin{aligned}\Big|\hat{H}^{N,j}(\underline{\xi}_s^N, x, t_N) - \underline{\hat{H}}^{N,j}(\underline{\xi}_s^N, x, t_N)\Big| &= \Big|E_{\underline{\xi}_s^N}[d^{N,j}(x, \xi_{t_N}^{N,\mathsf{vm}}) - d^{N,j}(x, \underline{\xi}_{t_N}^{N,\mathsf{vm}})]\Big| \\ &\le 2\|r\|E_{\underline{\xi}_s^N}\Big[\sum_{y\in\bar{\mathcal{N}}_N}(\xi_{t_N}^{N,\mathsf{vm}}(x+y) - \underline{\xi}_{t_N}^{N,\mathsf{vm}}(x+y))\Big] \\ &= 2\|r\|\sum_{y\in\bar{\mathcal{N}}_N}\hat{E}[\underline{\xi}_s^N(B_{t_N}^{N,x+y}) - \underline{\xi}_s^N(\underline{B}_{t_N}^{N,x+y})]. \end{aligned} \tag{10.30}$$

Now define

$$\mathcal{A}(\delta) = \{x \in S_N \cap \tilde{I} : \inf_{y\in(\tilde{I})^c}|x-y| \le \delta\} \text{ and } \mathcal{A}(\delta)' = \{x \in S_N \cap \tilde{I} : \inf_{y\in(\tilde{I})^c}|x-y| > \delta\}.$$

We will decompose the sum in (10.29) into sums over $x \in \mathcal{A}(t_N^{1/3})$ and $x \in \mathcal{A}(t_N^{1/3})'$. By (10.30),

$$\begin{aligned}&\frac{(\log N)^3}{N'}\sum_{x\in\mathcal{A}(t_N^{1/3})}|\hat{H}^{N,j}(\underline{\xi}_s^N, x, t_N) - \underline{\hat{H}}^{N,j}(\underline{\xi}_s^N, x, t_N)| \\ &\le 2\|r\|\frac{(\log N)^3}{N'}\sum_{x\in\mathcal{A}(t_N^{1/3})}\hat{E}\Big[\sum_{y\in\bar{\mathcal{N}}_N}\underline{\xi}_s^N(B_{t_N}^{N,x+y})\Big] \\ &= 2\|r\|\frac{(\log N)^3}{N'}\sum_{x\in\mathcal{A}(t_N^{1/3})}\sum_{w\in S_N}\underline{\xi}_s^N(w)\sum_{y\in\bar{\mathcal{N}}_N}\hat{P}(B_{t_N}^{N,y} = w-x) \\ &\le 2\|r\|\frac{(\log N)^3}{N'}\Big[\sum_{w\notin\mathcal{A}(2t_N^{1/3})}\underline{\xi}_s^N(w)\sum_{y\in\bar{\mathcal{N}}_N}\hat{P}(|B_{t_N}^{N,y}|_\infty > t_N^{1/3}) + |\bar{\mathcal{N}}|\sum_{w\in\mathcal{A}(2t_N^{1/3})}\underline{\xi}_s^N(w)\Big] \\ &\le C_{10.31}|\bar{\mathcal{N}}|(\log N)^3\Big[X_s^N(\mathbf{1})(N^{-1}+t_N)t_N^{-2/3} + X_s^N(\mathcal{A}(2t_N^{1/3}))\Big], \end{aligned} \tag{10.31}$$

for some constant $C_{10.31}$. In the last line we used $\underline{X}^N \le X^N$ and Chebychev's inequality, and in the next to last inequality we noted that for $w \notin \mathcal{A}(2t_N^{1/3})$ and $x \in \mathcal{A}(t_N)^{1/3}$, we must have $|w-x| > t_N^{1/3}$. To bound $E[X_s^N(\mathcal{A}(2t_N^{1/3}))]$ we will need a bound on the mean measure similar to Lemma 7.19, but now in terms of the L^1 norm of $\Psi : S_N \to R_+$, $\|\Psi\|_1 = \frac{1}{N}\sum_{x\in S_N}\Psi(x)$. The result we need is

$$E(X_s^N(\Psi)) \le X_0^N(P_s^N(\Psi)) + C_T X_0^N(\mathbf{1})(\log N)^3\|\Psi\|_1\Big(1 + \log\Big(1 \vee \frac{1}{\|\Psi\|_1}\Big)\Big) \quad \forall s \le T'.$$

This follows easily from (7.103) (with $c = 0$) and (7.2) as in the proof of Lemma 8.6 of [9]. Use this with $\Psi = 1\{\mathcal{A}(2t_N^{1/3})\}$ to take means in (10.31), recalling that X_0^N is supported on I to bound $X_0^N(P_s^N(\Psi))$, and obtain

$$\frac{(\log N)^3}{N}\int_0^{(t-t_N)^+}\sum_{x\in\mathcal{A}(t_N^{1/3})}E\Big[|\hat{H}^{N,j}(\underline{\xi}_s^N,x,t_N)-\underline{\hat{H}}^{N,j}(\underline{\xi}_s^N,x,t_N)|\Big]ds\le C_{10.32}(T')X_0^N(\mathbf{1})(\log N)^{-1/6}. \tag{10.32}$$

See the derivation of (141) in [9] for the details.

If $x\in\mathcal{A}(t_N^{1/3})'$ and $y\in\bar{\mathcal{N}}_N$, then $\underline{\xi}_s^N(B_{t_N}^{N,x+y})-\underline{\xi}_s^N(\underline{B}_{t_N}^{N,x+y})\neq 0$ implies that $B^{N,x+y}$ exits $\tilde{I}$ before t_N, and hence moves a distance exceeding $t_N^{1/3}$ from x by time t_N. Therefore (10.30) implies

$$\Big|\hat{H}^{N,j}(\underline{\xi}_s^N,x,t_N,)-\underline{\hat{H}}^{N,j}(\underline{\xi}_s^N,x,t_N)\Big|\le 2\|r\|\sum_{y\in\bar{\mathcal{N}}_N}\widehat{E}\Big[\underline{\xi}_s^N(B_{t_N}^{N,x+y})1\Big\{\sup_{u\le t_N}|B_u^{N,x+y}-x|>t_N^{1/3}\Big\}\Big].$$

With this bound, using Proposition 7.6(a), we have

$$\begin{aligned}
&\frac{(\log N)^3}{N'}\int_0^{t-t_N^+}\sum_{x\in\mathcal{A}(t_N^{1/3})'}E\Big[|\hat{H}^{N,j}(\xi_s^N,x,t_N)-\underline{\hat{H}}^{N,j}(\xi_s^N,x,t_N)|\Big]ds\\
&\quad\le\frac{2\|r\|(\log N)^3}{N'}\int_0^t E\Big[\sum_{w\in S_N}\sum_{x\in\mathcal{A}(t_N^{1/3})'}\sum_{y\in\bar{\mathcal{N}}_N}\underline{\xi}_s^N(w)\widehat{P}\Big(B_{t_N}^{N,y}=w-x,\sup_{u\le t_N}|B_u^{N,y}|>t_N^{1/3}\Big)\Big]ds\\
&\quad\le 2\|r\|(\log N)^3\sum_{y\in\bar{\mathcal{N}}_N}\widehat{P}(\sup_{u\le t_N}|B_u^{N,y}|>t_N^{1/3})\int_0^t E[\underline{X}_s^N(\mathbf{1})]ds\\
&\quad\le 2\|r\|C_{7.6}tX_0^N(\mathbf{1})(\log N)^3\sum_{y\in\bar{\mathcal{N}}_N}\widehat{P}(\sup_{u\le t_N}|B_u^{N,y}|^2>t_N^{2/3})\\
&\quad\le C_{10.33}(T')X_0^N(\mathbf{1})(\log N)^{-10/3},
\end{aligned} \tag{10.33}$$

where the weak L^1 inequality for nonnegative submartingales is used in the last. Use the above and (10.32) in (10.29) to obtain

$$|\mathcal{V}_{j,4}|\le C_{10.34}(T')(\log N)^{-1/6}X_0^N(1). \tag{10.34}$$

The inequalities (10.25), (10.26), (10.27), (10.28) and (10.34) imply (10.11), and we are done. □

Proof of Lemma 10.2. From Section 7.1 we may recall the bounding biased voter model $\bar{\xi}^N$ with rate function $\bar{c}^{N,b}$ as in (7.6), constructed as in (6.2). We then construct the killed version of $\bar{\xi}$, denoted $\underline{\bar{\xi}}$, using the same equation with rate function

$$\underline{\bar{c}}^{N,b}(x,\xi)=\bar{c}^{N,b}(x,\xi)1(x\in\tilde{I}).$$

We assume all these processes have the same initial state $\underline{\xi}_0^N=\xi_0^N$, supported on $\tilde{I}$, as our previous processes. In Section 7.1 we verified condition (6.3) for the rates of $(\xi^N,\bar{\xi}^N)$ and for the rates of $(\xi^{N,\mathsf{vm}},\bar{\xi}^N)$. The same condition is then immediate for the corresponding killed processes as one simply multiplies the required inequalities by $1(x\in\tilde{I})$. So we may now apply (6.4) for the killed processes to conclude that

$$\underline{\xi}^N\le\underline{\bar{\xi}}^N\quad\text{and}\quad\underline{\xi}^{N,\mathsf{vm}}\le\underline{\bar{\xi}}^N. \tag{10.35}$$

If $\underline{\bar{X}}_t = \frac{1}{N'}\sum_{x\in S_N}\underline{\bar{\xi}}^N_t(x)\delta_x$ and $c_0 = \|r\|(2+\underline{p}^{-1})$, then we may use the last display on p. 113 of [11] to conclude that

$$E[\underline{\bar{X}}^N_t(\mathbf{1})] = \underline{X}^N_0(\underline{P}^N_t(\mathbf{1})) + c_0\int_0^t \frac{(\log N)^3}{N'}\sum_{x\in S_N}\underline{P}^N_{t-s}(\mathbf{1})(x)(1-\underline{\bar{\xi}}^N_s(x)\sum_{y\in\mathcal{N}_N}\underline{\bar{\xi}}^N_s(x+y)\,ds. \tag{10.36}$$

To see this, first note that the setting in [11] is for a general class of voter model perturbations. This includes our biased voter model with (the notation is from [11]) $\delta_N\equiv 0$, and $\beta_N(\{y\}) = c_0(\log N)^3$ for $y\in\mathcal{N}_N$, while β_N is zero otherwise. The above formula then follows from p. 113 of [11]. Now take the difference of (10.36) with (10.4) and use (7.39) to conclude (recall (10.35))

$$\begin{aligned}
E[\underline{\bar{X}}_{t_N}(\mathbf{1}) - \underline{X}^N_{t_N}(\mathbf{1})] &= E\Big[\int_0^{t_N}\Big(c_0\Big\{\frac{(\log N)^3}{N'}\sum_{x\in S_N}\underline{P}^N_{t_N-s}(\mathbf{1})(x)(1-\underline{\bar{\xi}}^N_s(x))\sum_{y\in\mathcal{N}_N}\underline{\bar{\xi}}^N_s(x+y)\Big\} \\
&\qquad - \sum_{j=2}^{3} d^{N,j}(x,\underline{\xi}^N_s,\mathbf{1})\Big)\,ds\Big] \qquad (10.37)\\
&\le c_0\int_0^{t_N}|\bar{\mathcal{N}}|(\log N)^3 E[\bar{X}^N_s(\mathbf{1}) + X^N_s(\mathbf{1})]\,ds \\
&\le c_0|\bar{\mathcal{N}}|(\log N)^3\int_0^{t_N}[e^{c_0(\log N)^3 s}X^N_0(\mathbf{1}) + CX^N_0(\mathbf{1})]\,ds \\
&\le C(\log N)^{-16}X^N_0(\mathbf{1}), \qquad (10.38)
\end{aligned}$$

where Lemma 4.1 of [10]), and Proposition 7.6 (a) are used in the next to last line. The same reasoning (in fact it is simpler as there are no drift terms) gives

$$E[\underline{\bar{X}}_{t_N}(\mathbf{1}) - \underline{X}^{N,\mathsf{vm}}_{t_N}(\mathbf{1})] \le C(\log N)^{-16}X^N_0(\mathbf{1}). \tag{10.39}$$

Now argue exactly as in the proof of Lemma 7.2, using (10.38) and (10.39) in place of Lemma 7.1, to obtain the following killed version of Lemma 7.2 (recall now $p=19$ from the definition of t_N):

$$\begin{aligned}
&\text{There is a } C_{10.40} \text{ so that for } j=2,3, \text{ all } T'>0,\ s\in[0,T'] \text{ and all } \Phi\in C_b([0,T']\times\mathbb{R}^2),\\
&E_{\underline{\xi}^N_0}[|d^{N,j}(s,\underline{\xi}^N_{t_N},\Phi) - d^{N,j}(s,\underline{\xi}^{N,\mathsf{vm}}_{t_N},\Phi)|] \le C_{10.40}\|\Phi\|_\infty(\log N)^{-13}\underline{X}^N_0(\mathbf{1}). \qquad (10.40)
\end{aligned}$$

The proof of Lemma 10.2 is now completed using the Markov property just as in the derivation of Lemma 7.3. ☐

Proof of Corollary 1.17. By Theorem 4.9 and Theorem 1.15 we have the percolation condition (4.18). The required result now follows by a comparison to supercritical oriented percolation and Theorem 1.15 itself, as in the proof of survival in Proposition 5.3 of [11]. In fact, the derivation is now a bit easier as the uniform bounds proved there are not required. ☐

11 Appendix: The $|\mathcal{N}| = 8$ case of Lemma 3.3

Recall that for $|\mathcal{N}| = 8$,

$$M(k,j) = \sum_{\text{odd } i\le j\wedge k}\binom{j}{i}\binom{|8-j}{k-i}, \qquad 1\le k,j\le 8, \tag{11.1}$$

$$a_\ell = \left(\frac{\ell}{8}\right)^q, \qquad 1\le \ell\le 8. \tag{11.2}$$

Our goal is to verify that if $\boldsymbol{\alpha} = \mathbf{a}\mathbf{M}^{-1}$ then there exists a $q_0 < 1$ such that

$$\alpha_\ell(q) > 0 \; \forall \; q_0 < q < 1, \; 1 \le \ell \le 8. \tag{11.3}$$

The conclusion of Lemma 3.3 then follows as described in Section 3.1.

It is straightforward to check that $\mathbf{M}$ given by (11.1) is

$$\mathbf{M} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 7 & 12 & 15 & 16 & 15 & 12 & 7 & 0 \\ 21 & 30 & 31 & 28 & 25 & 26 & 35 & 56 \\ 35 & 40 & 35 & 32 & 35 & 40 & 35 & 0 \\ 35 & 30 & 25 & 28 & 31 & 26 & 21 & 56 \\ 21 & 12 & 13 & 16 & 13 & 12 & 21 & 0 \\ 7 & 2 & 5 & 4 & 3 & 6 & 1 & 8 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0. \end{bmatrix}$$

Using `maple`, we obtain

$$\mathbf{M}^{-1} = \begin{bmatrix} -\frac{3}{64} & -\frac{1}{32} & -\frac{1}{64} & 0 & \frac{1}{64} & \frac{1}{32} & \frac{3}{64} & \frac{1}{16} \\ -\frac{7}{64} & -\frac{1}{32} & \frac{1}{64} & \frac{1}{32} & \frac{1}{64} & -\frac{1}{32} & -\frac{7}{64} & -\frac{7}{32} \\ -\frac{7}{64} & \frac{1}{32} & \frac{3}{64} & 0 & -\frac{3}{64} & -\frac{1}{32} & \frac{7}{64} & \frac{7}{16} \\ 0 & \frac{5}{64} & 0 & -\frac{3}{64} & 0 & \frac{5}{64} & 0 & -\frac{35}{64} \\ \frac{7}{64} & \frac{1}{32} & -\frac{3}{64} & 0 & \frac{3}{64} & -\frac{1}{32} & -\frac{7}{64} & \frac{7}{16} \\ \frac{7}{64} & -\frac{1}{32} & -\frac{1}{64} & \frac{1}{32} & -\frac{1}{64} & -\frac{1}{32} & \frac{7}{64} & -\frac{7}{32} \\ \frac{3}{64} & -\frac{1}{32} & \frac{1}{64} & 0 & -\frac{1}{64} & \frac{1}{32} & -\frac{3}{64} & \frac{1}{16} \\ \frac{1}{128} & -\frac{1}{128} & \frac{1}{128} & -\frac{1}{128} & \frac{1}{128} & -\frac{1}{128} & \frac{1}{128} & -\frac{1}{128} \end{bmatrix},$$

and it is straightforward but tedious to check that this is indeed correct by verifying that the product of the above matrices is the identity matrix.

Given $\mathbf{M}^{-1}$ and $\mathbf{a}$ as above, $\boldsymbol{\alpha} = \mathbf{a}\mathbf{M}^{-1}$ is given by

$$\begin{aligned}
\alpha_1(q) &= -\frac{3\left(\frac{1}{8}\right)^q}{64} - \frac{7\left(\frac{1}{4}\right)^q}{64} - \frac{7\left(\frac{3}{8}\right)^q}{64} + \frac{7\left(\frac{5}{8}\right)^q}{64} + \frac{7\left(\frac{3}{4}\right)^q}{64} + \frac{3\left(\frac{7}{8}\right)^q}{64} + \frac{1}{128} \\
\alpha_2(q) &= -\frac{\left(\frac{1}{8}\right)^q}{32} - \frac{\left(\frac{1}{4}\right)^q}{32} + \frac{\left(\frac{3}{8}\right)^q}{32} + \frac{5\left(\frac{1}{2}\right)^q}{64} + \frac{\left(\frac{5}{8}\right)^q}{32} - \frac{\left(\frac{3}{4}\right)^q}{32} - \frac{\left(\frac{7}{8}\right)^q}{32} - \frac{1}{128} \\
\alpha_3(q) &= -\frac{\left(\frac{1}{8}\right)^q}{64} + \frac{\left(\frac{1}{4}\right)^q}{64} + \frac{3\left(\frac{3}{8}\right)^q}{64} - \frac{3\left(\frac{5}{8}\right)^q}{64} - \frac{\left(\frac{3}{4}\right)^q}{64} + \frac{\left(\frac{7}{8}\right)^q}{64} + \frac{1}{128} \\
\alpha_4(q) &= \frac{\left(\frac{1}{4}\right)^q}{32} - \frac{3\left(\frac{1}{2}\right)^q}{64} + \frac{\left(\frac{3}{4}\right)^q}{32} - \frac{1}{128} \\
\alpha_5(q) &= \frac{\left(\frac{1}{8}\right)^q}{64} + \frac{\left(\frac{1}{4}\right)^q}{64} - \frac{3\left(\frac{3}{8}\right)^q}{64} + \frac{3\left(\frac{5}{8}\right)^q}{64} - \frac{\left(\frac{3}{4}\right)^q}{64} - \frac{\left(\frac{7}{8}\right)^q}{64} + \frac{1}{128} \\
\alpha_6(q) &= \frac{\left(\frac{1}{8}\right)^q}{32} - \frac{\left(\frac{1}{4}\right)^q}{32} - \frac{\left(\frac{3}{8}\right)^q}{32} + \frac{5\left(\frac{1}{2}\right)^q}{64} - \frac{\left(\frac{5}{8}\right)^q}{32} - \frac{\left(\frac{3}{4}\right)^q}{32} + \frac{\left(\frac{7}{8}\right)^q}{32} - \frac{1}{128} \\
\alpha_7(q) &= \frac{3\left(\frac{1}{8}\right)^q}{64} - \frac{7\left(\frac{1}{4}\right)^q}{64} + \frac{7\left(\frac{3}{8}\right)^q}{64} - \frac{7\left(\frac{5}{8}\right)^q}{64} + \frac{7\left(\frac{3}{4}\right)^q}{64} - \frac{3\left(\frac{7}{8}\right)^q}{64} + \frac{1}{128} \\
\alpha_8(q) &= \frac{\left(\frac{1}{8}\right)^q}{16} - \frac{7\left(\frac{1}{4}\right)^q}{32} + \frac{7\left(\frac{3}{8}\right)^q}{16} - \frac{35\left(\frac{1}{2}\right)^q}{64} + \frac{7\left(\frac{5}{8}\right)^q}{16} - \frac{7\left(\frac{3}{4}\right)^q}{32} + \frac{\left(\frac{7}{8}\right)^q}{16} - \frac{1}{128}
\end{aligned}$$

By pairing off terms it is easy to see that $\alpha_1(q) \ge 1/128 > 0$ for all $0 \le q \le 1$. As described in Section 3.1, to prove (11.3) it remains only to verify (3.12), so we consider

$\alpha'_\ell(q), 2 \le \ell \le 8$. By differentiating, plugging in $q = 1$ and simplifying, we obtain

$$\begin{aligned}
\alpha'_2(1) &= \frac{3}{128}\log 2 - \frac{3}{256}\log 3 + \frac{5}{256}\log 5 - \frac{7}{256}\log 7\\
\alpha'_3(1) &= \frac{1}{64}\log 2 + \frac{3}{512}\log 3 - \frac{15}{512}\log 5 + \frac{7}{512}\log 7\\
\alpha'_4(1) &= -\frac{5}{128}\log 2 + \frac{3}{128}\log 3\\
\alpha'_5(1) &= \frac{1}{64}\log 2 - \frac{15}{512}\log 3 + \frac{15}{512}\log 5 - \frac{7}{512}\log 7\\
\alpha'_6(1) &= \frac{3}{128}\log 2 - \frac{9}{256}\log 3 - \frac{5}{256}\log 5 + \frac{7}{256}\log 7\\
\alpha'_7(1) &= \frac{5}{64}\log 2 + \frac{63}{512}\log 3 - \frac{35}{512}\log 5 - \frac{21}{512}\log 7\\
\alpha'_8(1) &= -\frac{101}{128}\log 2 + \frac{35}{128}\log 5 + \frac{7}{128}\log 7
\end{aligned}$$

We need only verify that each of the above are strictly negative. Resorting again to `maple`, the above derivatives can be written in the form

$$\begin{aligned}
\alpha'_2(1) &= -\frac{1}{256}\log\left(\frac{22235661}{200000}\right)\\
\alpha'_3(1) &= -\frac{1}{512}\log\left(\frac{30517578125}{5692329216}\right)\\
\alpha'_4(1) &= -\frac{1}{128}\log\left(\frac{32}{27}\right)\\
\alpha'_5(1) &= -\frac{1}{512}\log\left(\frac{11816941917501}{7812500000000}\right)\\
\alpha'_6(1) &= -\frac{1}{256}\log\left(\frac{61509375}{52706752}\right)\\
\alpha'_7(1) &= -\frac{1}{512}\log\left(\frac{16255824130589724722085520625114444091796875}{12584584288393115541569846261901038211563 52}\right)\\
\alpha'_8(1) &= -\frac{1}{128}\log\left(\frac{2535301200456458802993406410752}{2396825584582984447479248046875}\right).
\end{aligned}$$

This completes the proof of (3.12) for the case $|\mathcal{N}| = 8$, and so we are done.

References

[1] Abramiuk, A. and Sznajd-Weron, K. (2020) Generalized Independence in the q-Voter Model: How Do Parameters Influence the Phase Transition? *Entropy* **22**, 120, MR4072060.

[2] Abramiuk-Szurlej, A., Lipiecki, A., Pawlowski, J. and Sznajda-Weron, K.Discontinuous phase transitions in the q-voter model with generalized anticonformity on random graphs. *Sci. Reports* **11**, 17719, arXiv:2107.03338

[3] Abrams, D.M. and Strogatz, S. Modelling the dynamics of language death (2003) *Nature* **424**, 900.

[4] Agarwal, P., Simper, M. and Durrett, R. (2021) The q-voter model on the torus. *Electron. J. Probab.* **26** 1-33, MR4315513

[5] Cox, J.T., and Durrett, R. (1991) Nonlinear voter models. *Random Walks, Brownian Motion and Interacting Particle Systems*, R. Durrett and H. Kesten editors, 203-213, Birkhauser, MR1146446

[6] Cox, J.T., Durrett, R. and Perkins, E.A. (2000) Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* **18**, 185-234, MR1756003

[7] Cox, J.T., Durrett, R. and Perkins, E.A. (2013) Voter model perturbations and reaction diffusion equations. *Astérisque*, **349**, MR3075759

[8] Cox, J.T. and Griffeath, D. (1986) Diffusive clustering in the two dimensional voter model. Ann. Probab. **14**, 340-379, MR0832014

[9] Cox, J.T., Merle, M. and Perkins, E.A. (2010) Co-existence in a two-dimensional Lotka-Volterra model. *Electron. J. Probab.* **15**, 1190-1266, MR2678390

[10] Cox, J.T. and Perkins, E.A. (2005) Rescaled Lotka-Volterra models converge to super-Brownian motion. *Ann. Probab.* **33**, 904-947, MR2135308

[11] Cox, J.T. and Perkins, E.A. (2007) Survival and coexistence in stochastic spatial Lotka-Volterra models. *Prob. Theory Rel. Fields* **139**, 89-142 MR2322693

[12] Cox, J.T. and Perkins, E.A. (2008) Renormalization of the Two-dimensional Lotka-Volterra Model. *Ann. Appl. Prob.* **18**, 747-812, MR2399711

[13] Cox, J.T. and Perkins, E.A. (2014) A complete convergence theorem for voter model perturbations. *Ann. Appl. Prob.* **24**, 150-197, MR3161645

[14] Castellano, C. and Muñoz, M.A. and Pastor-Satorras, R. (2009) Nonlinear q-voter model. *Phys. Rev. E* **80**, 041129.

[15] Dawson, D.A. (1977) The critical measure diffusion process. *Z. Wahrsch. Verw. Gebiete* **40**, 125-145, MR0478374

[16] Durrett, R. (1995) Ten Lectures on Particle Systems, in *École d'Été de Probabilités de Saint Flour*, Lecture Notes Math. **1608**, 97-201, Springer-Verlag, Berlin, MR1383122

[17] Durrett, R., Liggett, T.M., and Zhang, Y. (2014) The contact process with fast voting. *Electron. J. Probab.* **19**, 1-19, MR3174840

[18] Ethier, S.N. and Kurtz, T.G. (1986) *Markov Processes, Characterization and Convergence.* John Wiley and Sons, New York, MR0838085

[19] Felsenstein, J. (1975) A Pain in the Torus: Some Difficulties with Models of Isolation by Distance. *American Naturalist* **109**, 359–368.

[20] Griffeath, D. (1979) Additive and cancellative interacting particle systems. *Springer Lecture Notes in Math.* **714**, MR0538077

[21] Handjani, S. (1999) The complete convergence theorem for coexistent voter models. *Ann. Probab.* **27**, 226-245, MR1681118

[22] Jedrzejewski, A. and Sznajd-Weron, K. (2022) Pair approximation for the q-voter models with quenched disorder on networks. *Phys. Rev. E* **105**, 064306, MR4455345

[23] Kallenberg, O. (1977) Stability of critical cluster fields. *Math. Nachr* **77**, 7-43, MR0443078

[24] Liggett, T.M. (1985). *Interacting Particle Systems*, Springer-Verlag, New York, MR2108619

[25] Liggett, T.M. (1999). *Stochastic interacting systems: contact, voter and exclusion processes*, Springer-Verlag, New York, MR1717346

[26] Moretti, P., Liu, S. and Castellano, C. and Pastor-Satorras, S. (2013) Mean-field analysis of the q-voter model on networks. *Stat. Phys.* **151**, 113-130, MR3045809

[27] Min, B. and San Miguel, M. (2017) Fragmentation transitions in a coevolving nonlinear voter model. *Sci. Reports* **7**, 12864, arXiv:1707.08111

[28] Nettle, D. (1999) Using Social Impact Theory to simulate language change. *Lingua* **108**, 95-117.

[29] Neuhauser, C. and Pacala, S.W. (1999) An explicitly spatial version of the Lotka-Volterra model with interspecific competition. *Ann. Appl. Probab.* **9**, 1226-1259, MR1728561

[30] Nowak, B., Ston, B. and Sznajd-Weron, K. (2020) Discontinuous phase transitions in the multi-state noisy q-voter model: quenched vs. annealed disorder. *Sci Rep* **11**, 6098.

[31] Sturm, A. and Swart, J. (2008) Voter models with heterozygosity selection. *Ann. Appl. Probab.* **18**, 59-99, MR2380891

[32] Vieira, A. and Anteneodo, C. (2018) Threshold q-voter model. *Phys. Rev. E* **97**, 052106, arXiv:1807.03661