The Annals of Applied Probability
2017, Vol. 27, No. 2, 1235–1288
DOI: 10.1214/16-AAP1230

EVOLVING VOTER MODEL ON DENSE RANDOM GRAPHS

BY RIDDHIPRATIM BASU[1] AND ALLAN SLY[2]

Stanford University and University of California, Berkeley

In this paper, we examine a variant of the voter model on a dynamically changing network where agents have the option of changing their friends rather than changing their opinions. We analyse, in the context of dense random graphs, two models considered in Durrett et al. [*Proc. Natl. Acad. Sci. USA* **109** (2012) 3682–3687]. When an edge with two agents holding different opinion is updated, with probability $\frac{\beta}{n}$, one agent performs a voter model step and changes its opinion to copy the other, and with probability $1-\frac{\beta}{n}$, the edge between them is broken and reconnected to a new agent chosen randomly from (i) the whole network (rewire-to-random model) or, (ii) the agents having the same opinion (rewire-to-same model). We rigorously establish in both the models, the time for this dynamics to terminate exhibits a phase transition in the model parameter β. For β sufficiently small, with high probability the network rapidly splits into two disconnected communities with opposing opinions, whereas for β large enough the dynamics runs for longer and the density of opinion changes significantly before the process stops. In the rewire-to-random model, we show that a positive fraction of both opinions survive with high probability.

1. Introduction. In recent years, a significant research effort in various fields, including biology, ecology, economics, sociology among others, has been concentrated on studying and modelling behaviour of large complex networks with many interacting agents. Different dynamics on large networks has been studied focussing on the structural impact of these dynamics on different models of networks. Some of the problems which received attention are consensus of opinion and polarisation, spread of epidemics, information cascades, etc. (see [3, 16]). In many real world networks, the evolution of the links in the network depend upon the states of the connecting agents and vice versa. The general class of network models that model this dependence are called adaptive or coevolutionary networks (see [7, 19]). As in the case of static networks, the problems of spread of information and epidemic, evolution of opinion and polarization into communities have been studied numerically and also using a variety of rigorous and partly nonrigorous methods ([6, 8, 12, 24, 25]; see also [4, 22] and references therein for more background). The problems we consider in this paper belong to this general class.

Received September 2015; revised April 2016.
[1]Supported by UC Berkeley Graduate Fellowship.
[2]Supported by an Alfred Sloan Fellowship and NSF Grants DMS-12-08338 and DMS-13-52013.
MSC2010 subject classifications. 60K35, 60K37, 05C80, 05C81.
Key words and phrases. Voter model, coevolutionary network, dense Erdős–Rényi graphs.

The voter model has classically been studied in the probability literature as an interacting particle system mainly on lattices [9, 14]. More recently, voter models have been studied in the context of general networks as a model for spread of opinion [20, 21]. In the classical voter model on a fixed graph, each vertex has one of the two prevalent opinions, neighbours interact at some fixed rate and one of the neighbours adopts the opinion of the other after the interaction. A simplified model of coevolution of network and opinion was introduced and studied using nonrigorous methods by Holme and Newman [10] where they try to model the property that an agent is less likely to interact with (remain connected with) another agent if their opinions do not match. Their model is similar to the classical voter model (with number of opinions proportional to the size of the network) but with the added feature that, whenever there is an interaction between two vertices (agents) with different opinion, with probability $\alpha \in (0, 1)$, one of the vertices breaks the link and connects to a different vertex of the same opinion, that is, the network connections evolve with time as well. Using finite size scaling, Holme and Newman conjectured a phase transition in α, where in the supercritical phase all the opinions eventually will have small number of followers, but in the subcritical phase a giant community holding the same opinion will emerge. This model and its further extensions were investigated in [11, 23].

Durrett et al. [4] studied two variants of this model on sparse random graphs using certain nonrigorous and numerical methods and formulated some conjectures about the asymptotic behaviour of the models as network size becomes large. They take the initial network to be a sparse Erdős–Rényi graph G with average degree greater than 1 and the two initial opinions distributed as product measure with density $u \in (0, \frac{1}{2}]$ of 1's. In each step, a uniformly chosen disagreeing edge is selected and a voter model step is performed with probability $1 - \alpha$ and a rewiring step with probability α. They consider two variants of reconnecting edge (i) rewire-to-random where the edge is connected to a randomly chosen vertex and (ii) rewire-to-same where the edge is connected to a random vertex of the same opinion. Based on numerical evidence and heuristics Durrett et al. conjecture in [4] that:

(1) Supercritical phase: In both variants, there exist $\alpha_c(u) \in (0, 1)$ (independent of u in the rewire-to-same model) such that for $\alpha > \alpha_c$, the process reaches an absorbing state in time $O(n)$ and the final fraction ρ of minority opinion is $\approx u$.

(2) Subcritical phase: In the *rewire-to-random* model, for $\alpha < \alpha_c(u)$, the time to absorption is order n^2 on average and at the absorption time, the density of the minority opinion ρ is bounded away from 0 and independent of u. In contrast, for the *rewire-to-same* model, $\rho \approx 0$, so one of the opinions takes over almost the whole network at the time of absorption.

We rigorously establish analogous results in the dense case but without establishing a sharp transition. We also prove that both opinions survive in *rewire-to-random* model, however, we cannot prove the contrasting result for the *rewire-to-same* model as is conjectured in [4]. Durrett et al. also formulate conjectures about finer behaviours of the evolving voter model along the path to absorption (see Conjecture 1 and Conjecture 2 in [4]). Further extensions of these models with different social dynamics and multiple possible opinions were considered in [15, 17, 18, 26].

1.1. *Main results*. In this paper, we study the dense version of the model from [4] where the initial graph is $G(n, 1/2)$ with density $\frac{1}{2}$ of both the opinions. It is easy to see that to obtain a nontrivial transition, we must renormalise the opinion update rate to $1-\alpha = \beta/n$ (this is due to the average degree being linear in n). In the sequel, whenever we say some event happens with high probability, it means that the event happens with probability tending to 1, as the number of vertices $n \to \infty$. Also by saying that an event occurs with exponentially high probability, we shall mean that the complement of the said event has probability that is exponentially small in n. Let τ denote the time to reach an absorbing state, that is, τ is the first time when there are no disagreeing edges in the graph (for the rewire-to-same model absorbing states are slightly different, see Section 1.2 below). For $\frac{1}{2} > \varepsilon > 0$, let $\tau_*(\varepsilon)$ be the first time that the fraction of the minority opinion reaches ε, that is, $\tau_*(\varepsilon) = \min\{t : N_*(t) \le \varepsilon n\}$, where $N_*(t)$ is the number of vertices holding the minority opinion at time t. Now we state our main theorems.

THEOREM 1. *Let $\frac{1}{2} > \varepsilon' > 0$ be given. For both variants of the model, there exist $0 < \beta_0 < \beta_*(\varepsilon') < \infty$ such that each of the following hold*:

(i) *For all $\beta < \beta_0$ and any $\eta > 0$, we have $\{\tau \le 6n^2, N_*(\tau) \ge (\frac{1}{2} - \eta)n\}$ holds with high probability as $n \to \infty$.*

(ii) *For all $\beta > \beta_*(\varepsilon')$, we have that $\tau_*(\varepsilon') \le \tau$ with high probability as $n \to \infty$ and*

$$\lim_{c \downarrow 0} \liminf_{n} \mathbb{P}\big[\tau > cn^3\big] = 1.$$

The following theorem asserts that for the rewire-to-random model the fraction of the minority opinion vertices is bounded away from 0 at the absorption time.

THEOREM 2. *Let $\beta > 0$ be fixed. For the rewire-to-random model there exists $\varepsilon_* = \varepsilon_*(\beta) > 0$ such that $\tau < \tau_*(\varepsilon_*)$ with high probability.*

1.2. *Formal model definitions*. Now we describe formally the models we consider in this paper. Let n be a fixed positive integer. Let V be a fixed set with $|V| = n$. Let $V^{(2)}$ denote the set of all unordered pairs in V, we shall call elements of $V^{(2)}$ as *bonds*. Also let $\mathcal{E}$ be a fixed set with $|\mathcal{E}| = N$. We consider discrete time Markov chains $\{G(t)\}_{t \ge 0}$ taking values in

$$\{\{0,1\}^V, (V^{(2)})^{\mathcal{E}}\},$$

that is, for each t, $G(t)$ is a multi-graph on the vertex set V with labelled edges coming from the set $\mathcal{E}$ (each edge in $\mathcal{E}$ is placed at one of the bonds); each vertex has one of the two opinions 0 and 1. The following notation will be used throughout this paper.

The opinion of a vertex v at time t shall be denoted by $v(t)$. The vector of opinions of vertices in $G(t)$ shall be denoted by $V(t)$. We shall denote by $N_0(t)$ and $N_1(t)$ the number of 0s and 1s in $V(t)$, respectively. Let $N_*(t) = \min\{N_0(t), N_1(t)\}$. For $v \in V$, let $C_v(t)$ denote the set of all vertices in V which have the same opinion as v in $G(t)$. By $\tilde{G}(t) = (V, E(t))$, we denote the underlying graph of $G(t)$, that is, $\tilde{G}(t)$ is a (multi)graph with vertex set V and edge set $E(t)$. Often, when there is no scope of confusion we shall use $G(t)$ instead of $\tilde{G}(t)$ to denote the same. Notice that we are allowing multi-edges but not self loops, that is, at a time t, a bond $(u, v) \in V^{(2)}$ may be connected by more than one edge, but there are no edges connecting v to itself. For an edge connecting the bond (u, v) in $G(t)$, we shall call it *disagreeing* if $u(t) \neq v(t)$ and *agreeing* otherwise.

Initial condition: To simplify matters, we only consider the following initial condition. We take $\tilde{G}(0)$ is distributed as $G(n, \frac{1}{2})$, that is, each bond contains 0 edge with probability $\frac{1}{2}$ and 1 edge with probability $\frac{1}{2}$ independent of each other. Also, let $\{v(0)\}_{v \in V}$ be i.i.d. $\mathrm{Ber}(\frac{1}{2})$. Also denote the set of the labelled edges $\mathcal{E} = \{e_0, e_1, \dots, e_N\}$.

Transition probabilities: We describe the one step evolution of the two variants of the Markov chain as follows.

Let $G(t)$ be the state of the chain at time t. Let $Z(t) = \mathrm{Ber}(\frac{\beta}{n})$ be independent of $G(t)$. If $Z(t) = 1$, we obtain $G(t+1)$ from $G(t)$ by taking a *voter model step*, and if $Z(t) = 0$ then we obtain $G(t+1)$ from $G(t)$ by taking a *rewiring step*. We call $\beta > 0$ the *relabelling rate*; this is a parameter of the model.

Let $\mathcal{E}^{\times}(t) \subseteq \mathcal{E}$ denote the set of edges that are disagreeing in $G(t)$. Choose one edge e from $\mathcal{E}^{\times}(t)$ uniformly at random. Let (u, v) be the bond which this edge connects in $G(t)$. Choose one vertex randomly among u and v, say u. The vertex u as above will be called the *root* of a relabelling/rewiring move.

The voter model step (relabelling step): If $Z(t) = 1$, then u adopts the opinion of v, that is, $G(t+1)$ is obtained from $G(t)$ by taking $\tilde{G}(t+1) = \tilde{G}(t)$, $v'(t+1) = v'(t)$ for all $v' \in V \setminus \{u\}$ and $u(t+1) = v(t+1)$.

The rewiring step: If $Z(t) = 0$, the two chains we consider evolve differently.

Rewire-to-random model: In this model, we choose a vertex v' uniformly at random from $V \setminus \{u\}$. We obtain $G(t+1)$ from $G(t)$ by taking $V(t+1) = V(t)$ and $E(t+1)$ is obtained from $E(t)$ by removing the edge e from the bond (u, v) and adding it to the bond (u, v').

Rewire-to-same model: In this model, we choose a vertex v' uniformly at random from $C_u(t) \setminus \{u\}$. We obtain $G(t+1)$ from $G(t)$ by taking $V(t+1) = V(t)$ and $E(t+1)$ is obtained from $E(t)$ by removing the edge e from the bond (u, v) and adding it to the bond (u, v').

We make the following basic observation characterising the absorbing states.

OBSERVATION 1.1. *On finite networks both the chains are absorbing. For the rewire-to-random model, the only absorbing states are those which corresponds to the graph having no disagreeing edges, that is, either one opinion has taken over all the vertices, or the graph is split into disconnected communities, where all the vertices in a community has the same opinion. For the rewire-to-same model, the absorbing states are those that either have no disagreeing edges, or those in which one of the opinions are held by only one vertex.*

Notice that the number of edges is conserved in each step of the chain, that is, we have that $|E(t)| = |\mathcal{E}|$ for all t. Also observe that even though we have labelled edges, this fact does not affect the behaviour of the model at all. The edges are labelled simply because it will be convenient while constructing a coupling of this chain with another process which we shall use.

One of the main questions we are interested in for both the models described above is the asymptotics of absorption time as a function of β as $n \to \infty$, and whether it exhibits a phase transition in β or not. Notice that if $\beta = 0$, then we have only rewiring moves and the absorption time is $\Theta(n^2)$, that is, the graph splits immediately into two communities having different opinions. We investigate whether similar phenomenon occurs if $\beta > 0$ is sufficiently small. In the other extreme, if the rewiring moves are much rarer compared to the relabelling moves (i.e., $\beta \gg 1$) one might expect the model to behave similarly as the voter model on a static graph, where the minority opinion density will become very small before reaching an absorbing state, and the absorption time will be at least $\Theta(n^3)$. This is established in Theorem 1. A related quantity of interest is the fraction of the minority opinion vertices when the process reaches an absorbing state. For β sufficiently large does the minority opinion persist with a positive fraction? Theorem 2 provides the answer for the rewire-to-random model.

1.3. *Outline of the proof.* We prove parts (i) and (ii) of Theorem 1 separately. The arguments are similar for the rewire-to-random model and the rewire-to-same model. We provide details only for the rewire-to-random model while pointing out the differences for the rewire-to-same model. To prove part (i), we essentially show that before the density of either opinion changes, a rewiring move is likely to decrease the number of disagreeing edges. Using a martingale argument, we show that, by time $\Theta(n^2)$ (by which time the opinion densities cannot change significantly), the number of disagreeing edges decay to 0.

Most of the work goes into proving part (ii) of Theorem 1. We show that for $\beta = \beta(\varepsilon')$ sufficiently large, with high probability the graph $G(t)$ remains close enough to an Erdős–Rényi graph, in a sense made precise later, as long as the minority opinion density does not drop below ε'. To this end, we define a number of stopping times detecting when $G(t)$ deviates too much from an Erdős–Rényi graph for the first time with respect to certain different properties, and roughly show that all those stopping times are with high probability at least as large as

$\tau_*(\varepsilon')$. The properties we need to consider are vertex degrees, the Cheeger constant and edge-multiplicities.

Corresponding to each of the properties we consider, we define two stopping times, one with a stronger threshold and the other with a weaker threshold. We show that provided none of the weaker thresholds have been reached, the opinions quickly mix to an approximate product measure which guarantees that the properties of interest are sufficiently mean reverting for our purposes and with high probability the stronger thresholds are also not reached.

For the proof of Theorem 2, we show that for a fixed β, there exists a sufficiently small but positive ε_*, such that once the minority opinion reaches ε_*, the typical vertices having minority opinions start losing disagreeing edges at a higher rate than it gains them, and eventually all the disagreeing edges are lost before the minority opinion density can substantially decrease further.

Organisation of the paper: The rest of the paper is organised as follows. In Section 2, we prove part (i) of Theorem 1 for the rewire-to-random model. Most of the work in this paper goes toward the proof of part (ii) of Theorem 1 for the rewire-to-random model, which spans Section 3, Section 4 and Section 5. In Section 3, we define all the stopping times that we need to use. In Section 4, we show that, if by time t the graph does not reach any of the stronger stopping times, then for β sufficiently large the graph does not reach any of the weaker stopping times by time $t+\delta n^2$ with high probability, where δ is a small constant. That the graph is also unlikely to reach any of the strong stopping times by time $t+\delta n^2$ as long as the minority opinion density does not become too small is shown in Section 5. Together these complete the proof of Theorem 1, part (ii). Theorem 2 is proved in Section 6. In Section 7, we point out the significant adaptations to the argument that are necessary to prove Theorem 1 for the rewire-to-same model. We finish with the discussion of some open problems in Section 8.

2. Fast polarization for small β.

In this section, we prove part (i) of Theorem 1 for the *rewire-to-random model* with relabelling rate β. First, we make the following definitions. Let $D_{\max}(t)$ denote the maximum degree of a vertex in $G(t)$. The degree of a vertex is defined as the number of edges incident to it, and not the number of bonds containing edges. Also, let $X_t = |\mathcal{E}^{\times}(t)|$ denote the number of disagreeing edges at time t. Consider the following stopping times. Let $\tau_1 = \min\{t : D_{\max}(t) \geq 8n\}$ and let $\tau_2 = \tau_*(\frac{1}{3})$, that is, $\tau_2 = \min\{t : N_*(t) = \min\{N_0(t), N_1(t)\} \leq \frac{n}{3}\}$. Define $\tau_0 = \tau \wedge \tau_1 \wedge \tau_2$. We have the following lemmas.

LEMMA 2.1. *There exists $\beta_0 > 0$, such that for all $\beta < \beta_0$, we have $\tau_0 \leq 6n^2$ with high probability.*

PROOF. Let $\mathcal{F}_t$ denote the filtration generated by the process up to time t. Observe that whenever an edge is rewired, X_t either remains the same or decreases by 1. Conditional on $\mathcal{F}_t$, the chance that X_t is decreased by a rewiring move is at least $\frac{N_*(t)-1}{n-1}$. Also notice that a relabelling move, that is, a voter model step can increase X_t by at most $D_{\max}(t)$. Hence, we have for $\lambda > 0$,

$$\begin{aligned}&\mathbb{E}\big(e^{\frac{\lambda X_{t+1}}{n}} \mid \mathcal{F}_t\big)\\ &\quad\le e^{\frac{\lambda X_t}{n}}\left(\left(1-\frac{\beta}{n}\right)\left(1+\frac{N_*(t)-1}{n-1}\big(e^{-\frac{\lambda}{n}}-1\big)\right)+\frac{\beta}{n}e^{\frac{\lambda D_{\max}(t)}{n}}\right).\end{aligned} \tag{2.1}$$

Now for large n on the event $\{t < \tau_0\}$ we have $\frac{N_*(t)-1}{n-1} \ge \frac{1}{4}$. Take $\lambda > 0$ sufficiently small so that $e^{8\lambda} \le 1 + 9\lambda$ and $e^{-\frac{\lambda}{n}} - 1 \le -\frac{\lambda}{2n}$. Then on $\{t < \tau_0\}$, we have for $\beta < \frac{1}{400}$,

$$\begin{aligned}\mathbb{E}\big(e^{\frac{\lambda X_{t+1}}{n}} \mid \mathcal{F}_t\big) &\le e^{\frac{\lambda X_t}{n}}\left(\left(1-\frac{\beta}{n}\right)\left(1+\frac{1}{4}\big(e^{-\frac{\lambda}{n}}-1\big)\right)+\frac{\beta}{n}e^{8\lambda}\right)\\ &\le e^{\frac{\lambda X_t}{n}}\left(\left(1-\frac{\beta}{n}\right)\left(1-\frac{\lambda}{8n}\right)+\frac{\beta}{n}(1+9\lambda)\right)\\ &\le e^{\frac{\lambda X_t}{n}}\left(1-\frac{\lambda}{10n}\right).\end{aligned} \tag{2.2}$$

It follows from above that

$$\begin{aligned}\mathbb{P}[\tau_0 > t \mid \mathcal{F}_0] &\le \mathbb{E}\big[e^{\frac{\lambda X_t}{n}} 1_{\{\tau_0 > t\}} \mid \mathcal{F}_0\big]\\ &\le e^{\frac{\lambda X_0}{n}} e^{-\frac{\lambda t}{10n}}\\ &\le e^{\frac{\lambda n}{2}} e^{-\frac{\lambda t}{10n}}\end{aligned} \tag{2.3}$$

since $X_0 \le \frac{n^2}{2}$. Hence, we have

$$\mathbb{P}\big[\tau_0 > 6n^2\big] \le e^{-\frac{\lambda n}{10}}. \qquad \square$$

LEMMA 2.2. *We have $\tau_1 > 6n^2$ with high probability.*

PROOF. It is easy to see that the *rewire-to-random* dynamics can be implemented in the following way. Without loss of generality, let $V = [n]$. And let $\mathbb{W} = \{W_i\}_{i\ge 1}$ be a sequence of i.i.d. random variables with each W_i being uniformly distributed over $\{1, 2, \dots, n\}$. Let us define $L_0 = 0$, and define L_i for $i > 0$ recursively as follows. Let, v_i be the *root* of the ith rewiring move. Then we define $L_i = \min\{j > L_{i-1} : W_j \ne v_i\}$. Then in the ith rewiring move, we add the edge to the bond (v_i, W_{L_i}). The algorithm can be described as follows. For each rewiring move, start inspecting the list $\mathbb{W}$ from the first previously uninspected element up to the first time you find a vertex which is not equal to the root of the current rewiring. Rewire the edge to this vertex. Clearly, in this way the chosen vertex is uniform among all vertices other than v_i, and hence this is indeed an implementation of the *rewire-to-random* dynamics.

Now observe that $L_{i+1} - L_i$ are i.i.d. $\mathrm{Geom}(\frac{n-1}{n})$ variables. It follows by a large deviation estimate that $L_{6n^2} < \frac{13n^2}{2}$ with exponentially high probability. Now for $v \in V$, let

$$N(v) = \#\left\{i \leq \frac{13n^2}{2} : W_i = v\right\}.$$

Clearly, $N(v)$ is distributed as $\mathrm{Bin}(\frac{13n^2}{2}, \frac{1}{n})$, and a Chernoff bound implies

$$\mathbb{P}\bigl[N(v) \geq 7n\bigr] \leq e^{-n/78}.$$

Hence, noting that $D_{\max}(0) \leq n$, we have using a union bound over all the vertices

$$\mathbb{P}\bigl[\tau_1 \leq 6n^2\bigr] \leq ne^{-n/78} + \mathbb{P}\left[L_{6n^2} > \frac{13n^2}{2}\right]. \tag{2.4}$$

This completes the proof of the lemma. □

LEMMA 2.3. *There exists $\beta_0 > 0$, such that for all $\beta < \beta_0$, we have $\tau_2 \geq 6n^2 \wedge \tau$ with high probability.*

PROOF. For $t \geq 1$, it is easy to see that $RL(t)$, the number of relabelling moves up to time t, is stochastically dominated by a $\mathrm{Bin}(t, \frac{\beta}{n})$ variable. On $\{t < \tau\}$, we have that $N_0(t) - N_0(0)$ is distributed as $Z_{RL(t)}$ where $\{Z_i\}_{i\geq 0}$ is a simple symmetric random walk on $\mathbb{Z}$ started from 0. Using a union bound, it follows that

$$\begin{aligned}
&\mathbb{P}\bigl[\tau_2 < 6n^2 \wedge \tau\bigr] \\
&\quad \leq \mathbb{P}\left[N_*(0) \leq \frac{2n}{5}\right] + \mathbb{P}\bigl[RL\bigl(6n^2\bigr) > 12\beta n\bigr] + \mathbb{P}\left[\max_{i \leq 12\beta n} |Z_i| \geq \frac{3n}{20}\right].
\end{aligned}$$

By choosing β_0 sufficiently small, the last term in the above inequality is 0 for all $\beta < \beta_0$. Noticing that $\mathbb{P}[N_*(0) < \frac{2n}{5}] = 2\mathbb{P}[\mathrm{Bin}(n, \frac{1}{2}) < \frac{2n}{5}]$ and using Hoeffding inequality to bound the first term and a Chernoff bound on the second term yields

$$\mathbb{P}\bigl[\tau_2 < 6n^2 \wedge \tau\bigr] \leq 2e^{-2n/25} + e^{-2\beta n}. \tag{2.5}$$

This completes the proof of the lemma. □

Now we are ready to prove Theorem 1(i).

PROOF OF THEOREM 1(i). From Lemma 2.1, Lemma 2.2 and Lemma 2.3, it follows that with high probability we have $\{\tau_0 \leq 6n^2, \tau_1 \wedge \tau_2 \geq 6n^2 \wedge \tau\}$. It follows that, $\tau \leq 6n^2$ with high probability. The second part of the theorem follows from noting that using a random walk estimate as in Lemma 2.3, we see that for each $\eta > 0$, the probability that the density of the minority opinion drops below $\frac{1}{2} - \eta$ within $6n^2$ steps tends to 0 as $n \to \infty$. This completes the proof of the theorem. □

3. High relabelling rate case: Stopping times.

3.1. *A time change*: rewire-to-random-* *dynamics*. For the proof of Theorem 1(ii), we shall consider a time changed variant of *rewire-to-random* dynamics, which we call *rewire-to-random-** model. This model is same as the *rewire-to-random* model, except that now at time $(t+1)$, instead of choosing a disagreeing edge at random, we choose an edge at random from $G(t)$. If the edge is not disagreeing, then we do nothing. It is clear that *rewite-to-random-** model is a slowed down version of *rewire-to-random* model. It is also clear that if we prove Theorem 1(ii) for the *rewire-to-random-** dynamics, then it will imply the same theorem for the *rewire-to-random* dynamics.

Assumption on the initial condition: For this section and the next two, we shall always assume that $G(0)$ satisfies the following conditions:

(i) $|E(0)|$, the number of edges in $G(0)$ is in $[\frac{n^2}{4} - n^{3/2}, \frac{n^2}{4} + n^{3/2}]$.
(ii) $\#\{v \in V : v(0) = 0\} \in [\frac{n}{2} - n^{3/4}, \frac{n}{2} + n^{3/4}]$.

Since both the events hold with probability $1 - o(1)$, this assumption does not affect any of our results.

Now we move toward proving Theorem 1(ii). Let us fix $\varepsilon' > \varepsilon > 0$ for the rest of this paper. Let $\tau_* = \tau_*(\varepsilon)$, that is,

$$\tau_* = \min\{t : \min(N_0(t), N_1(t)) \le \varepsilon n\}.$$

Parameters: Now we define a number of stopping times. In the definition of these stopping times and the proofs that follow, we use a number of parameters that need to satisfy the following constraints. For a fixed ε, our parameters satisfy the following inequalities. We choose C_1 sufficiently large depending on ε, the exact functional dependence not being of consequence. We then choose ε_2 sufficiently small such that $\varepsilon_2 < \varepsilon^2/1000$ and $\varepsilon_2 \log \frac{2C_1}{\varepsilon_2} < \frac{\varepsilon}{16}$. Then we choose $\varepsilon_3 < \varepsilon_2^2/1000$ and $\varepsilon_7 < \varepsilon_3^2/1000$. Fixing these parameters, we choose $C_2 = 2$, $\varepsilon_4 < \frac{1}{4\log 10}$, $\delta < \frac{\varepsilon_3}{10{,}000 C_2} \wedge 10^{-10}$. We choose $0 < \varepsilon_{14} < \varepsilon^2/100 \wedge \frac{(\frac{\varepsilon}{8} - 2\varepsilon_2 \log \frac{2C_1}{\varepsilon_2})^2}{4C_2}$. After fixing all these parameters, C is chosen sufficiently large depending on these. There are also many other parameters used in the proofs which are chosen either sufficiently small or large depending on other parameters, again where the functional dependence is not of importance to us. Finally, β is taken sufficiently large depending on all the parameters used. Also we shall always take n sufficiently large depending on everything else.

• *Stopping times for large cuts*:

Let S and T be two disjoints subsets of V with $S \cup T = V$. We denote by $N_{ST}(t)$ the number of edges in $G(t)$ with one endpoint in S and another endpoint in T. Define $N_{SS}(t)$ and $N_{TT}(t)$ similarly. Also let $N(t) = N$ denote the total number of edges in $G(t)$. Define

$$K_{ST}(t) = \left(\frac{N_{SS}(t) - \frac{1}{4}|S|^2}{N(t)}\right)^2 + \left(\frac{N_{TT}(t) - \frac{1}{4}|T|^2}{N(t)}\right)^2.$$

Set

$$L(t) = \max_{S,T:\min(|S|,|T|)\geq \varepsilon_2 n} K_{ST}(t)$$

and

$$L'(t) = \max_{S,T:\min(|S|,|T|)\geq \varepsilon_2 n} \left|\frac{N_{ST}(t) - \frac{1}{2}|S||T|}{N(t)}\right| \vee \left|\frac{N_{SS}(t) - \frac{1}{4}|S|^2}{N(t)}\right|.$$

Now the two stopping times are defined as follows:

- The stronger stopping time: $\tau_2 = \min\{t : L(t) \geq \varepsilon_3^2\}$.
- The weaker stopping time: $\tau_2' = \min\{t : L'(t) \geq 2\varepsilon_3\}$.

Notice that in the definitions above, the quantities we center by are expected value of N_{ST} (resp., N_{SS}, etc.) if the underlying graph were $G(n, \frac{1}{2})$. Thus, these stopping times control the "distance" of $G(t)$ from an Erdős–Rényi $G(n, \frac{1}{2})$ in terms of the number of edges across a cut (S, T) if neither of the sets S and T is too small.

• *Stopping times for individual edge multiplicities*:

For $u, v \in V$, let $M_{uv}(t)$ denote the number of edges in the bond (u, v) in $G(t)$. Let $M(t) = \max_{u\neq v} M_{uv}(t)$. If the time t is clear from the context we shall drop it from the above notation. Now the two stopping times are defined as follows:

- The stronger stopping time: $\tau_3 = \min\{t : M(t) \geq \varepsilon_4 \log n\}$.
- The weaker stopping time: $\tau_3' = \min\{t : M(t) \geq 2\varepsilon_4 \log n\}$.

• *Stopping times for balanced vertices*:

Let us call a vertex v ϵ-balanced in $G(t)$ if for all k, $\#\{u \in V : M_{uv}(t) \geq k\} \leq \epsilon 10^{-k} n$. We define the two stopping times as follows:

- The stronger stopping time: $\tau_4 = \min\{t : \exists v \in V \text{not} C_1\text{-balanced in} G(t)\}$.
- The weaker stopping time: $\tau_4' = \min\{t : \exists v \in V \text{not} 2C_1\text{-balanced in} G(t)\}$.

• *Stopping times for maximum and minimum degrees*:

Let $D_{\max}(t)$ and $D_{\min}(t)$ denote the maximum and minimum degree in $G(t)$, respectively. The stopping times are defined as follows:

- The stronger stopping time: $\tau_5 = \min\{t : D_{\max}(t) > (1 - \frac{\varepsilon}{2})n \text{or} D_{\min}(t) < \frac{\varepsilon n}{2}\}$.
- The weaker stopping time: $\tau_5' = \min\{t : D_{\max}(t) > C_2 n \text{or} D_{\min}(t) < \frac{\varepsilon n}{4}\}$.

The following lemma is immediate from the definitions and we omit the proof.

LEMMA 3.1. *For each $i = 2, 3, 4, 5$, we have $\tau_i' \leq \tau_i$.*

One should interpret the above stopping times as follows. On $t < \tau_2 \wedge \tau_3 \wedge \tau_4 \wedge \tau_5$, the graph $G(t)$ is sufficiently "close" to the random graph $G(0)$ that we stared with. The same is true on $t < \tau_2' \wedge \tau_3' \wedge \tau_4' \wedge \tau_5'$ but the thresholds in the definition of being "close" is now somewhat weaker.

Finally, we define $\tau_0 = \tau_* \wedge \tau_2 \wedge \tau_3 \wedge \tau_4 \wedge \tau_5$ and $\tau_0' = \tau_* \wedge \tau_2' \wedge \tau_3' \wedge \tau_4' \wedge \tau_5'$.

Part (ii) of Theorem 1 will follow from the next theorem which shows that the process cannot reach the stopping time τ_0 too much ahead of reaching the stopping time τ_*.

THEOREM 3.2. *There exist $\beta_* = \beta_*(\varepsilon)$ such that for all $\beta > \beta_*$, we have for the rewire-to-random-* model $\tau_0 \geq \tau_* - n^2$ with high probability.*

We shall prove Theorem 3.2 over the next two sections. Before that we show how this implies part (ii) of Theorem 1 for the *rewire-to-random-** model.

PROOF OF THEOREM 1(II). Notice that, it follows from a random walk estimate that $\tau_* \geq \tau_*(\varepsilon') + n^2$ with high probability. On $\{\tau_0 \geq \tau_*(\varepsilon')\}$, we have that $\tau_*(\varepsilon') - 1 < \tau_0$. And hence, in particular, $\tau_*(\varepsilon') - 1 < \tau_2$. Let S be the set of all vertices with the minority opinion at time $\tau_* - 1$. Since $\varepsilon_2 < \varepsilon < \varepsilon'$ and $\tau_2' \geq \tau_2$, we have that $N_{ST}(\tau_* - 1) \geq \frac{1}{2}|S||T| - 2\varepsilon_3 N(t) > 0$, since $2\varepsilon_3 < \varepsilon_2(1-\varepsilon_2)$. It then follows that $\tau \geq \tau_*(\varepsilon')$ for the *rewire-to-random-** dynamics. Since the *rewire-to-random-** dynamics is merely a time changed version of the *rewire-to-random* dynamics, the first statement in Theorem 1(ii) follows.

To prove the second statement, observe the following. For $t \geq 0$, Let $S(t)$ denote the number of voter model steps up to time t. A random walk estimate as before establishes that for $\beta > \beta_*$ (given by Theorem 3.2),

$$\lim_{c \downarrow 0} \liminf_{n} \mathbb{P}\bigl[S\bigl(\tau_*(\varepsilon)\bigr) > 2c\beta n^2\bigr] = 1.$$

Since the relative frequency of voter model steps is $\frac{\beta}{n}$, it follows that for any fixed c and β $\mathbb{P}[S(\tau_*(\varepsilon)) > 2c\beta n^2, \tau_*(\varepsilon) < cn^3] = o(1)$ as $n \to \infty$. The result now follows from Theorem 3.2. $\square$

Before proceeding with the proof of Theorem 3.2, further let us recall the overall strategy. Given that the process has not reached the any of the stronger stopping times τ_2, τ_3, τ_4 and τ_5 as well as τ_* by some time t (i.e., on $\{t < \tau_0\}$) we shall show that it is unlikely that by time $t + \delta n^2$ the process reaches the stopping time τ_0' without reaching the stopping time τ_*. Once we have established that, we establish that given that the process has not reached τ_0 by time t, it is unlikely at time $t + \delta n^2$ that the process will violate any of the conditions of the stronger stopping time τ_0 without reaching the weaker stopping time τ_0' before that time. The proof will then be completed by taking a union bound over times $k\delta n^2$, $k = 1, 2, \dots$.

We start by proving the following lemma which establishes the connection between the evolving voter model dynamics and our stopping times that we shall exploit extensively.

LEMMA 3.3. *Consider a single walker performing the following continuous time random walk on $G(t)$. Let each directed edge ring at rate $\frac{\beta}{2n}$. Whenever an edge rings if the walker is at the starting point of the edge, it moves along the edge. Let $\lambda(G(t))$ denote the spectral gap of this Markov chain. Then there exists $\frac{\varepsilon^2}{100} > \varepsilon_{14} > 0$, such that we have, on $\{t < \tau_2' \wedge \tau_3' \wedge \tau_4' \wedge \tau_5'\}$, $\lambda(G(t)) \geq \beta\varepsilon_{14}$.*

PROOF. Let $h(G(t))$ denote the Cheeger constant of the corresponding random walk. We have

$$h(G(t)) := \min_{S,T:S\cup T=V, S\cap T=\varnothing, |S|\leq n/2} \frac{N_{ST}(t)\beta}{2|S|n}.$$

Now on $\{t \leq \tau_2'\}$, if $|S| \geq \varepsilon_2 n$, $N_{ST}(t) \geq \frac{1}{2}|S||T| - 2\varepsilon_3 N(t) \geq \frac{1}{2}|S||T| - \varepsilon_3 n^2$, and hence

$$\frac{N_{ST}(t)\beta}{2|S|n} \geq \beta\left(\frac{1}{8} - \frac{\varepsilon_3}{2\varepsilon_2}\right) \geq 2\beta\sqrt{C_2}\sqrt{\varepsilon_{14}}$$

provided

$$\varepsilon_{14} \leq \frac{(\frac{1}{8} - \frac{\varepsilon_3}{2\varepsilon_2})^2}{4C_2}.$$

On $\{t \leq \tau_4' \wedge \tau_5'\}$, we have

$$N_{SS}(t) + N_{ST}(t) \geq \frac{\varepsilon|S|n}{4}. \tag{3.1}$$

Now observe that by the definition of balanced vertices, on $\{t \leq \tau_4' \wedge \tau_5'\}$ for any vertex v, the total number of edges from v to any subset of at most $\varepsilon_2 n$ vertices is bounded above by

$$\sum_{k\geq k_0} 2C_1 k 10^{-k} n, \tag{3.2}$$

where k_0 is the smallest integer satisfying

$$\sum_{k\geq k_0} 2C_1 10^{-k} n \leq \varepsilon_2 n.$$

Indeed, we get the above bound by considering the $\varepsilon_2 n$ bonds adjacent to v that contain maximum number of edges. It is easy to see that $k_0 \approx \log\frac{20C_1}{9\varepsilon_2}$ and the quantity in (3.2) is bounded by $2n\varepsilon_2 \log\frac{2C_1}{\varepsilon_2}$. It follows that if $|S| \leq \varepsilon_2 n$ we have

$$N_{SS}(t) \leq 2|S|n\varepsilon_2 \log\frac{2C_1}{\varepsilon_2}. \tag{3.3}$$

Combining (3.1) and (3.3), it follows that

$$\frac{N_{ST}(t)\beta}{2|S|n} \geq \beta\left(\frac{\varepsilon}{8} - 2\varepsilon_2 \log \frac{2C_1}{\varepsilon_2}\right) \geq 2\beta\sqrt{C_2}\sqrt{\varepsilon_{14}}$$

provided

$$\varepsilon_{14} \leq \frac{(\frac{\varepsilon}{8} - 2\varepsilon_2 \log \frac{2C_1}{\varepsilon_2})^2}{4C_2}.$$

Now it follows that on $\{t \leq \tau_2' \wedge \tau_4' \wedge \tau_5'\}$,

$$h(G(t)) \geq 2\beta\sqrt{C_2}\sqrt{\varepsilon_{14}}.$$

Now, using Cheeger inequality (Theorem 13.14 of [13], see [5], equation (1.9), for the variant used here) we get

$$\lambda(G(t)) \geq \frac{h(G(t))^2}{2\frac{\beta D_{\max}(t)}{n}} \geq \frac{h(G(t))^2}{2\beta C_2} \geq \beta\varepsilon_{14}$$

which completes the proof of the lemma. □

4. Estimates for the weak stopping times. In this section, we show that provided the process has not reached any of the stronger stopping times by time t, that is, $t < \tau_0$, then within a small number of steps (δn^2 in number), the process is unlikely to reach any of the weaker stopping times, that is, $t + \delta n^2 < \tau_0'$ with high probability. The general idea is that by time δn^2, there are not enough rewiring steps to change the graph substantially. We start with the following proposition which controls the fraction of minority opinion vertices in the time interval $\{t+1, t+2, \dots, t+\delta n^2\}$.

PROPOSITION 4.1. *We have that* $\mathbb{P}(\tau_*(\frac{4\varepsilon}{5}) \leq t + \delta n^2 \mid \mathcal{F}_t, t < \tau_0) \leq e^{-cn}$ *where* $c = c(\delta, \beta, \varepsilon) > 0$.

PROOF. It follows from a Chernoff bound that the probability that there are more than $2\delta\beta n$ many relabelling steps in $[t+1, t+\delta n^2]$ is exponentially small in n. Note that the number of vertices of a certain opinion does a simple symmetric random walk in absorbed at 0 or n in the *rewire-to-random* dynamics. The lemma now follows from a random walk estimate by observing that *rewire-to-random-** dynamics is slower than rewire-to-random dynamics. □

A remark is in order here. Observe that in the above proposition the condition $t < \tau_0$ is not necessary; we could have replaced it with the condition $t < \tau_*$. There will be a number of such occurrences in this and a few in the following section, where in the statement of the results we would assume $t < \tau_0$ or $t < \tau_0'$, although a weaker condition might suffice for the proofs of those particular results. We shall keep be the generic estimates in the statements, as it will be necessary in putting together estimates for different stopping times. However, in some of the significant cases we shall point out the extent to which the assumptions can be relaxed.

The next proposition considers the weaker stopping times for the large cuts. Note that the assumption $t < \tau_0$ can be replaced by $t < \tau_2$ in the statement below.

PROPOSITION 4.2. *On the event* $\{t < \tau_0\}$, $t + \delta n^2 < \tau_2'$.

PROOF. Clearly, for all S, T that make a partition of V and any t' with $t \le t' \le t + \delta n^2$, $|N_{ST}(t') - N_{ST}(t)| \vee |N_{SS}(t') - N_{SS}(t)| \le \delta n^2$. It follows from definitions that if $\delta < \frac{\varepsilon_3}{100}$, then $t + \delta n^2 \le \tau_2'$. $\square$

The next proposition shows the weaker degree estimates continue to hold until time $t + \delta n^2$ with high probability if $t < \tau_0$ (as a matter of fact $t < \tau_5$ suffices).

PROPOSITION 4.3. *We have* $\mathbb{P}(t + \delta n^2 \ge \tau_5' \mid \mathcal{F}_t, t < \tau_0) \le e^{-cn}$ *for some constant* $c > 0$.

PROOF. Condition on $\mathcal{F}_t$. For any fixed vertex v, the number of times in $[t + 1, t + \delta n^2]$ an edge is rewired to v is stochastically dominated by a $\mathrm{Bin}(\delta n^2, \frac{1}{n-1})$ variable. By Chernoff's inequality and a union bound, it follows that the probability that any vertex gets more than $2\delta n$ edges is exponentially small in n. It follows that with exponentially high probability $\max_{t' \in [t+1, t+\delta n^2]} D_{\max}(t') < C_2 n$ provided $C_2 > (1 - \frac{\varepsilon}{2}) + 2\delta$.

For the lower bound, we do the following. Fix a vertex v. By hypothesis at time t, the degree of v, denoted $D_v(t)$ is at least $\varepsilon n/2$. Consider the process $\{D_v(t') : t' \in [t, t + \delta n^2]\}$. Let $R_v(t')$ denote the number of edges v loses up to time t' while D_v is at most $\varepsilon n/2$. Clearly, if $R_v(t + \delta n^2) \le \varepsilon n/4$ then $D_v(t + \delta n^2) > \varepsilon n/4$. Now, at each time t', conditioned on everything up to time $t' - 1$, the chance that R_v increases by one is bounded above by $\frac{\varepsilon n/2}{n^2/5}$ (since one of the at most $\varepsilon n/2$ edges adjacent to v must ring out of at least $\frac{n^2}{5}$ edges). It follows that

$$\mathbb{P}\big(R_v\big(t + \delta n^2\big) \ge \varepsilon n/4\big) \le \mathbb{P}\left(\mathrm{Bin}\left(\delta n^2, \frac{10\varepsilon}{4n}\right) \ge \varepsilon n/4\right)$$

which in turn is exponentially small in n by a Chernoff bound provided $12\delta < 1$. Taking a union bound over all the vertices v completes the proof of the lemma. $\square$

Now we prove a similar statement for individual edge-multiplicities.

PROPOSITION 4.4. *We have* $\mathbb{P}(t + \delta n^2 \ge \tau_3' \mid \mathcal{F}_t, t < \tau_0) \le \frac{1}{n^{20}}$.

PROOF. Condition on $\{\mathcal{F}_t, t < \tau_0\}$. For every bond (u, v) in $V^{(2)}$, let A_{uv} denote the event that $D_u(t') + D_v(t') \le 2C_2 n$ for all $t' \in [t+1, t+\delta n^2]$ where $D_w(t')$ denotes the degree of the vertex w in $G(t')$. It follows from Lemma 4.3 that A_{uv} holds with exponentially high probability. Let R_{uv} denote the number of edges added to the bond (u, v) in time $[t+1, t+\delta n^2]$. On A_{uv} we have

$$R_{uv} \preceq \mathrm{Bin}\left(\delta n^2, \frac{10C_2}{n^2}\right),$$

where $\preceq$ denotes stochastic domination. Indeed, observe that on A_{uv} at each time $t' \in [t+1, t+\delta n^2]$ the chance that a new edge gets added to the bond (u, v) is at most $\frac{10C_2}{n^2}$ conditioned on everything till time $t' - 1$ (out of at least $\frac{2n^2}{9}$ edges in total, one out of the at most $2C_2 n$ incident to u or v must ring at step t for this, and the chance that it is rewired to the correct vertex is $\frac{1}{n-1}$). By a Chernoff bound again, provided n is sufficiently large we get that

$$\mathbb{P}\left(\mathrm{Bin}\left(\delta n^2, \frac{10C_2}{n^2}\right) \ge \varepsilon_4 \log n\right) \le \frac{1}{n^{23}}.$$

Taking a union bound over all bonds completes the proof of the lemma. $\square$

Note that in the above proposition we could have replaced the condition $t < \tau_0$ by $t < \tau_3 \wedge \tau_5$.

Finally, we work out the estimates for the number of multi-edges incident to a given vertex.

PROPOSITION 4.5. *We have* $\mathbb{P}(t + \delta n^2 \ge \tau_4' \mid \mathcal{F}_t, t < \tau_0) \le \frac{1}{n^{18}}$.

Fix $v \in V$ and $0 < k < 2\varepsilon_4 \log n$. Let $N(\ell \to k)$ denote the number of bonds containing v that had ℓ edges at time t and that gained at least $(k - \ell)$ edges during time $[t+1, t+\delta n^2]$. Clearly,

$$\max_{t' \in [t+1, t+\delta n^2]} \#\{u : M_{uv}(t') \ge k\} \le C_1 10^{-k} n + \sum_{\ell=0}^{k-1} N(\ell \to k). \tag{4.1}$$

Therefore, we need to control the quantities $N(\ell \to k)$ for $\ell = 0, 1, \dots, k-1$. We prove the following lemma.

LEMMA 4.6. *Fix $v \in V$ and $0 < k < 2\varepsilon_4 \log n$. For $\ell < k$ and $N(\ell \to k)$ as above, we have*

$$\mathbb{P}\left(N(\ell \to k) > 25^{-(k-\ell)} C_1 10^{-k} n \mid \mathcal{F}_t, t < \tau_0\right) \le e^{-\frac{C_1 \sqrt{n}}{100}}.$$

Condition on $\mathcal{F}_t, t < \tau_0$. Fix v, ℓ and k as in the statement of the lemma. As in the proof of Proposition 4.4, we shall ignore without loss of generality the event with exponentially small probability that for some $t' \in [t+1, t+\delta n^2]$ there exists $u, v \in V$ such that $D_u(t') + D_v(t') > 2C_2 n$. Let $u_1, u_2, \dots, u_D$ be the set of vertices in V such that $M_{vu_i}(t) = \ell$. Without loss of generality, we can assume $D = C_1 10^{-\ell} n$. Let T_i denote the number of edges gained by the bond (v, u_i) in $[t+1, t+\delta n^2]$. We construct a family of random variables $(Y_1, Y_2, \dots, Y_D)$ which jointly stochastically dominates $(T_1, T_2, \dots, T_D)$, denoted $(T_1, T_2, \dots, T_D) \preceq (Y_1, Y_2, \dots, Y_D)$, that is, for any $(a_1, a_2, \dots, a_D)$ we have

$$\mathbb{P}[T_i \geq a_i \ \forall i] \leq \mathbb{P}[Y_i \geq a_i \ \forall i].$$

LEMMA 4.7. *Consider the following "balls and bins" model. Start with D empty urns. For each $i = \{1, 2, \dots, \delta n^2\}$, at the ith step with probability $\frac{12C_2 D}{n^2}$ we choose one of urns uniformly and put a ball in it. Let $(Y_1, Y_2, \dots, Y_D)$ denote the vector of the number of balls in the urns after δn^2 steps. Then $(T_1, T_2, \dots, T_D) \preceq (Y_1, Y_2, \dots, Y_D)$.*

PROOF. Observe that, at each step in the voter model the conditional probability of a new edge being added to one of the bonds (v, u_i) is at most $\frac{12C_2}{n^2}$ (as shown in Lemma 4.4), which is equal to the chance that an edge is added to the ith urn in the balls and bins model at any step. It follows that there is a coupling between the voter model and the balls and bin model such that if the bond (v, u_i) gains an edge at time $t+h$ then the ith bin gains a ball at step h. The lemma follows. □

Notice that $(Y_1, Y_2, \dots, Y_D)$ above is not a collection of independent variables. However, to prove Lemma 4.6 we shall stochastically dominate $(Y_1, Y_2, \dots, Y_D)$ by a collection of independent binomials conditioned on a high probability event. We have the following lemma.

LEMMA 4.8. *Let $(Y_1, Y_2, \dots, Y_D)$ be as in Lemma 4.7. Let $(Y'_1, Y'_2, \dots, Y'_D)$ denote a collection of i.i.d.* $\mathrm{Bin}(24\delta C_2 D, \frac{2}{D})$ *variables. Then the conditional joint law of $(Y_1, Y_2, \dots, Y_D)$ given $\sum_{i=1}^{D} Y_i \leq 24\delta C_2 D$ is stochastically dominated by the conditional joint law of $(Y'_1, Y'_2, \dots, Y'_D)$ given $\sum_{i=1}^{D} Y'_i \geq 24\delta C_2 D$.*

PROOF. Fix $\mathcal{X} \in \mathbb{N}$. One way to sample the joint distribution of $(Y_1, Y_2, \dots, Y_D)$ conditioned on $\sum Y_i = \mathcal{X}$ is the following. For any $p \in [0, 1]$, consider throwing i.i.d. $Y'_i \sim \mathrm{Bin}(\mathcal{X}, p)$ balls into each of D bins. Condition on the event $\sum_i Y'_i \geq \mathcal{X}$, choose uniformly $\mathcal{X}$ of the balls and color them black. Now let Y''_i denote the number of black balls in the ith bin. Clearly $(Y_1, Y_2, \dots, Y_D)$ conditioned on $\sum Y_i = \mathcal{X} \in \mathbb{N}$ had the same distribution as $(Y''_1, Y''_2, \dots, Y''_D)$ conditioned on $\sum_i Y'_i \geq \mathcal{X}$. Since $Y'_i \geq Y''_i$ for all i, we get the lemma by specializing to $\mathcal{X} = 24C_2 D$, $p = \frac{2}{D}$ and noting that for $\mathcal{X}_1 \leq \mathcal{X}_2$, the distribution of $(Y_1, Y_2, \dots, Y_D)$ conditioned on $\sum Y_i = \mathcal{X}_1$ is stochastically dominated by the distribution of $(Y_1, Y_2, \dots, Y_D)$ conditioned on $\sum Y_i = \mathcal{X}_2$. □

Now we are ready to prove Lemma 4.6.

PROOF OF LEMMA 4.6. Let $\mathcal{C}$ (resp., $\mathcal{C}'$) denote the event that $\sum_{i=1}^{D} Y_i \leq 24\delta C_2 D$ (resp., $\sum_{i=1}^{D} Y_i' \geq 24\delta C_2 D$). Further, let Z_i (resp., Z_i') denote the indicator of $Y_i \geq (k-\ell)$ [resp., $Y_i' \geq (k-\ell)$]. It follows from Lemma 4.8 that

$$\begin{aligned}\mathbb{P}\left[\sum_{i=1}^{D} Z_i \geq 25^{-(k-\ell)} D\right] &\leq \mathbb{P}[\mathcal{C}^c] + \mathbb{P}\left[\sum_{i=1}^{D} Z_i \geq 25^{-(k-\ell)} D \mid \mathcal{C}\right] \\ &\leq e^{-4\delta C_2 D} + \mathbb{P}\left[\sum_{i=1}^{D} Z_i' \geq 25^{-(k-\ell)} D \mid \mathcal{C}'\right] \\ &\leq e^{-4\delta C_2 D} + 2 \times \mathbb{P}\big[\mathrm{Bin}\big(D, q(k-\ell)\big) \geq 25^{-(k-\ell)} D\big],\end{aligned}$$

where $q(k-\ell) = \mathbb{P}(\mathrm{Bin}(24\delta C_2 D, \frac{2}{D}) \geq (k-\ell))$. In the above equation, we have used Chernoff bounds to deduce $\mathbb{P}[\mathcal{C}^c] \leq e^{-4\delta C_2 D}$ and $\mathbb{P}[\mathcal{C}'] \geq \frac{1}{2}$. Using another Chernoff bound, we get

$$q(k-\ell) \leq e^{-\frac{1}{2}(k-\ell)\log(\frac{k-\ell}{48C_2\delta})} \leq 30^{-(k-\ell)}$$

provided $1 > 48{,}000 C_2 \delta$.

Finally, as observed before, without loss of generality we assume $D = C_1 10^{-\ell} n$ and by Lemma 4.7 we have $N(\ell \to k) \preceq \sum Z_i$. Using another Chenoff bound, we get

$$\begin{aligned}\mathbb{P}\big(N(\ell \to k) \geq (25)^{-(k-\ell)} C_1 10^{-k} n\big) &\leq 2e^{-\frac{C_1\sqrt{n}}{75}} + e^{-4\delta C_1 C_2 \sqrt{n}/100} \\ &\leq e^{-\frac{C_1 10^{-k} n}{100}}\end{aligned}$$

since $k < 2\varepsilon_4 \log n$ provided $2\varepsilon_4 \log 10 < \frac{1}{2}$. $\square$

Observe from the proof that we could have replaced the condition $t < \tau_0$ in this lemma by the weaker condition $t < \tau_4 \wedge \tau_5$. The same is true of Proposition 4.5, which we are now ready to prove.

PROOF OF PROPOSITION 4.5. Condition on $\{\mathcal{F}_t, t < \tau_0\}$. Using Lemma 4.6, a union bound over $\ell \in [0, k-1]$ and (4.1) we get that for a fixed v

$$\mathbb{P}\Big[\max_{t' \in [t+1, t+\delta n^2]} \#\{u : M_{uv}(t') \geq k\} > 2C_1 10^{-k} n\Big] \leq \frac{1}{n^{25}}.$$

A union bound over all $k \in [1, 2\varepsilon_4 \log n]$, Lemma 4.4 and another union bound over all vertices v complete the proof of the proposition. $\square$

All the propositions in this section together imply the following theorem.

THEOREM 4.9. *For the rewire-to-random-* model, we have* $\mathbb{P}(t+\delta n^2 < \tau_*, t+\delta n^2 \geq \tau_0' \mid \mathcal{F}_t, t<\tau_0) \leq \frac{1}{n^{17}}$.

5. Estimates for strong stopping times. Our goal in this section is to prove that if by time t the process does not reach any of the strong stopping times, then it is also unlikely that any of the strong stopping times will be hit by time $t+\delta n^2$ unless the minority opinion density drops below ε. We shall prove this by separate analysis of each of the stopping times. In the heart of the analysis, in each case, is some estimate on how the opinions of the vertices get mixed in a short (compared to δn^2) time for β sufficiently large, which we prove by constructing a coupling of a random walk on the graph $G(t)$ with the evolving voter model dynamics.

5.1. *The coupling construction.* The dual relationship between the random walk and the (continuous-time) voter model on a fixed graph H with vertex set $V(H)$ is well known. The distribution of the voter model $X(t)=\{X_u(t): u \in V(H)\}$ started from $X(0)$ can be constructed by running coalescing random walks for time t from each vertex and setting the opinion $X_u(t)$ of the vertex u at time t to be the value of $X(0)$ at the location of the walker started from u (cf. Section 1.7, [3]). We prove that an analogue of that result holds in our set-up. We want to say that if t is small so that graph has not changed sufficiently in the meantime, then the two distributions are not far apart. Now we formally describe the coupling.

An equivalent implementation of the rewire-to-random-* *dynamics starting at time* $t+1$: Let us condition on $\mathcal{F}_t$. Let $N=N(t)$ be the number of edges in $G(t)$. Let the set of all labelled edges be $\{e_1, e_2, \ldots, e_N\}$. For the purpose of this subsection, we shall assume that these edges are directed, that is, the edges have two identifiable ends e_i^+ and e_i^-. When an edge is placed in a bond, we think of it as e_i^+ being placed at one vertex of the bond and e_i^- being placed at the other. Suppose e_i is placed in the bond (u,v) and e_i^+ is placed at u and e_i^- placed at v. Then if a rewiring move rewires e with root u to the vertex w, then e_i^- is placed at w. Consider the two independent sequences $\mathbb{RW}=\{RW_i\}_{i\geq 1}$ and $\mathbb{RL}=\{RL_i\}_{i\in\mathbb{Z}}$ where each RL_i are chosen independently and uniformly from the set $E^*=\{e_1^+, e_1^-, e_2^+, e_2^-, \ldots, e_N^+, e_N^-\}$. The sequence $\mathbb{RW}$ is also an i.i.d. sequence where each, $RW_i=(e^*, v)$, where e^* is picked uniformly from $E^*=\{e_1^+, e_1^-, e_2^+, e_2^-, \ldots, e_N^+, e_N^-\}$, and v is a uniform random vertex v picked from V independently of e^*. Also let $\{Z_i\}_{i\geq 1}$ be a sequence of i.i.d. $\mathrm{Ber}(\frac{\beta}{n})$ variables.

We construct the equivalent formulation of the process as follows. At time $t+i$, if $Z_i=0$, then choose the first uninspected element from $\mathbb{RW}$, let that element be (e_j^+, v). If the edge e_j is not disagreeing in $G(t+i-1)$, then do nothing. Otherwise, we try to rewire the edge e_j to v, with the root as the vertex having e_j^+. If this rewiring is not legal, [i.e., e_j^+ is already placed at v in $G(t+i-1)$], then we choose the next element from $\mathbb{RW}$ and repeat the process until a successful rewiring. If $Z_i = 1$, then we choose the first unused element from $\mathbb{RL}$, let that element be e_j^+. If the edge e_j is not disagreeing in $G(t+i-1)$ then do nothing. Otherwise, we change the opinion of the vertex containing e_j^+. Notice that $\mathbb{RL}$ is a bi-infinite sequence but we start inspecting the elements starting from RL_1. It is clear from our construction that this indeed is an equivalent implementation of the *rewire-to-random-** dynamics. Let us run this dynamics starting with $G(t)$ up to $\frac{Cn^2}{\beta}$ steps. Let σ and ω be the number of elements of $\mathbb{RL}$ and $\mathbb{RW}$ that gets inspected in the process. Notice that σ is independent of the sequences $\mathbb{RL}$ and $\mathbb{RW}$.

Coupling with continuous time random walks: Now consider the following continuous time random walk on $G(t)$. Each directed edge rings at rate $\frac{\beta}{2n}$. When a directed edge rings a walker at the starting point of the edge moves along the edge. Consider the process where we start with one walker at each vertex of some arbitrary subset $W \subseteq V$, and each walker independently performs the random walk described above. We consider the following coupling between this process and the evolving voter model process described above. To start with each of the random walks are of type A. Now choose T_i i.i.d. $\exp(2N)$. At the ith step, wait time T_i. If $Z_{Cn^2/\beta+1-i} = 0$, then do nothing. If $Z_{Cn^2/\beta+1-i} = 1$, then look at $RL_{\sigma+1-k(i)}$ where $k(i) = \#\{j \in [Cn^2/\beta + 1 - i, Cn^2/\beta] : Z_j = 1\}$. If there is any walker of type A at the starting point of that edge, then that walker takes a step along that edge. If any walker of type A takes a step to a vertex where there is already one or more walkers, then all the walkers become of type B. Type B walkers do a random walk having the same waiting time distributions but using independent randomness. It is clear that the random walks are independent. Also since σ is independent of $\mathbb{RL}$, the random walks also have correct marginals. So this is indeed a coupling as we claimed.

Let $\mathbb{T} = \sum_{i=1}^{Cn^2/\beta} T_i$. Let us make the following definitions. For $v \in V$, let

$$E^*(v) = \{e^* \in E^* : (e^*, v) = RW_i \text{ for some } i < \omega\}.$$

DEFINITION 5.1. Consider the random walks described above. At time s, we call the walker starting at v_0 *happy* if the following conditions are satisfied:

1. It is of type A at time s.
2. None of the edges the walker has traversed have been rewired in the voter model process.
3. Let the path traversed by the walker be $\{v_0, v_1, \dots, v_k\}$, with the times of jumps being $0 = T_0^* < T_1^* < T_2^* < \cdots < T_k^* \le s$. Then for each ℓ, and each $e^* \in E^*(v_\ell)$ there was no ring in e^* in $[T_\ell^*, T_{\ell+1}^*]$.

The following lemma records the most basic useful fact about this coupling construction.

LEMMA 5.2. *Consider the coupling described above. At time* $\mathbb{T}$, *if a walker starting from* v *is happy, then the opinion of the position of that walker at time* $\mathbb{T}$ *is the same as the opinion of* v *in the* rewire-to-random-* *dynamics at time* $t + \frac{Cn^2}{\beta}$.

PROOF. This follows from the definition of the coupling. $\square$

LEMMA 5.3. *For the coupling described as above, let* $\mathbb{T} = \sum_{i=1}^{\frac{Cn^2}{\beta}} T_i$ *be as defined above. Then, for any* $\kappa_3 > 0$, *with exponentially high probability* $\frac{2C-\kappa_3}{\beta} \leq \mathbb{T} \leq \frac{2C+\kappa_3}{\beta}$.

PROOF. The result follows from a large deviation estimate for sum of independent exponential variables. $\square$

5.1.1. *Mixing time of the individual random walks.* Now consider a continuous time random walk on $G(t)$ whose transitions are given as follows. The walker at vertex v moves to a vertex u at rate $\frac{\beta M_{uv}(t)}{n}$. Clearly, this describes the random walk performed by each of the individual walkers described above. It is easy to see that the stationary measure for this random walk is the uniform measure on the vertices of $G(t)$. We show that if $\{t < \tau_0'\}$, that is, if $G(t)$ is sufficiently close to $G(0)$ in the relevant metrics then the walk mixes in $O(\frac{1}{\beta})$ time, that is, the distribution of the position of the walk becomes close to a uniform distribution.

For completeness, we recall the definitions of several metrics between probability distributions that we will use to quantify the closeness in distribution.

DEFINITION 5.4. For any two distributions μ and ν on a finite probability space Ω, define the total variation distance between μ and ν by

$$\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A \in \mathcal{P}} \mu(A) - \nu(A) = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

The L^1 and (relative) L^2-distance between μ and ν are defined respectively by

$$\|\mu - \nu\|_1 := \sum_{x \in \Omega} |\mu(x) - \nu(x)|;$$

$$\|\mu - \nu\|_2 = \left(\sum_{x \in \Omega} \frac{(\mu(x) - \nu(x))^2}{\nu(x)} \right)^{1/2}.$$

It is a basic fact that

$$2\|\mu - \nu\|_{\mathrm{TV}} = \|\mu - \nu\|_1 \leq \|\mu - \nu\|_2.$$

Also for random variable X, Y with distribution μ and ν, respectively, we shall use $\|X-Y\|_{\mathrm{TV}}$, $\|X-Y\|_1$ or $\|X-Y\|_2$ to refer to the corresponding distance between the respective distributions of X and Y.

The following proposition is the basic mixing time result on $G(t)$ that we shall need.

PROPOSITION 5.5. *Consider the continuous time random walk on $G(t)$ described above starting from an arbitrary vertex v. Let $Y(t+t')$ denote the position of the walk at time t'. Also let U be a uniformly chosen vertex from V. Then for sufficiently large C, we have that on $\{t<\tau_0'\}$, for all $t'\geq \frac{C}{\beta}$,*

$$\|Y(t+t')-U\|_{\mathrm{TV}}\leq e^{-\frac{\sqrt{C}}{1000}}.$$

PROOF. We know from Lemma 3.3 that on $\{t<\tau_0'\}$, $\lambda:=\lambda(G(t))\geq \beta\varepsilon_{14}$. Now let T_1 be the time of the first jump of the walker initially at v. Then we know by L^2 contraction lemma (Lemma 3.26 in [1])

$$\|Y(t+T_1+t^*)-U\|_2\leq e^{-t^*\lambda}\|Y(t+T_1)-U\|_2.$$

Now we know that on $\{t<\tau_0'\}$,

$$\begin{aligned}\|Y(t+T_1)-U\|_2^2&\leq \sum_{k=1}^{\infty}\frac{9k^2}{n^2}2C_1 10^{-k}n^2\\ &\leq 18C_1\sum_{k=1}^{\infty}k^2 10^{-k}\\ &\leq 50C_1.\end{aligned}$$

Now observe that for $t'\geq \frac{C}{\beta}$,

$$\begin{aligned}\|Y(t+t')-U\|_1&\leq 2\|Y(t+T_1+t'/2)-U\|_1+\mathbb{P}(T_1\geq t'/2)\\ &\leq 2e^{-\frac{t'\lambda}{2}}\|Y(t+T_1+t'/2)-U\|_2+\mathbb{P}(T_1\geq t'/2)\\ &\leq 10\sqrt{2C_1}e^{-t'\beta\varepsilon_{14}/2}+e^{-\frac{C\varepsilon}{16}}\\ &\leq e^{-C\varepsilon_{14}/20}+e^{-\frac{3C\varepsilon}{40}},\end{aligned}$$

where we have used the relationships between $\|\cdot\|_{\mathrm{TV}}$, $\|\cdot\|_1$ and $\|\cdot\|_2$. The lemma follows by taking C sufficiently large. $\square$

5.1.2. *Properties of the coupling.* In this subsection, we shall establish a few useful properties of the coupling described above. We start with establishing some basic estimates for a collection of independent random walks having the same law as the ones in the coupling. Notice that in the next five lemmas (Lemma 5.6–Lemma 5.10) one can replace the condition $t < \tau_0'$ by the weaker condition $t < \tau_4' \wedge \tau_5'$.

LEMMA 5.6. *Let $\kappa > 0$ be fixed. Let $v_1, v_2, \dots, v_{\varepsilon_{13}n}$ be given vertices in V. From each of the vertices, we run independent discrete time simple random walks in $G(t)$ up to $20C$ steps. Then, on $\{t < \tau_0'\}$, we have for ε_{13} small enough, with exponentially high probability there exists at least $(1-\kappa)\varepsilon_{13}n$ vertices among these such that the paths of random walks started from these vertices do not intersect.*

PROOF. Set $V^* = \{v_1, v_2, \dots, v_{\varepsilon_{13}n}\}$. Let D_1 be the set of vertices v_i such that v_i is hit by the random walk started from some v_j, $j \neq i$ and let D_2 be the set of vertices v_i such that the random walk started from v_i intersects a random walk started from v_j for some $j < i$. Clearly, for $i, j \in V^* \setminus (D_1 \cup D_2)$, random walks started from v_i and v_j do not intersect. Hence, it is sufficient to control the sizes of D_1 and D_2 only.

Condition on $\{\mathcal{F}_t, t < \tau_0'\}$. For $i = 1, 2, \dots, \varepsilon_{13}n$, and $j = 1, 2, \dots, 4C$, let Z_{ij} denote the indicator of the event that the random walk started from i hits the set $V^* = \{v_1, v_2, \dots, v_{\varepsilon_{13}n}\}$ in step j. Let $\mathcal{F}_{i,j}$ denote the filtration generated by the random walk paths of the walks started from $v_1, v_2, \dots, v_{i-1}$ and the first $(j-1)$ steps of the random walk started from v_i. Now notice that on $\{t < \tau_0'\}$, for any vertex v, the number of edges from v to V^* in $G(t)$ is at most

$$\begin{aligned}\sum_{j:\sum_{k\geq j} 2C_1 10^{-k} \leq \varepsilon_{13}} 2C_1 j 10^{-j} n &\leq \sum_{j \geq \log(\frac{4C_1}{\varepsilon_{13}})} 2C_1 j 10^{-j} n \\ &\leq 25\varepsilon_{13} \log\left(\frac{4C_1}{\varepsilon_{13}}\right) n \\ &\leq \frac{\varepsilon \kappa n}{400C}\end{aligned}$$

for ε_{13} sufficiently small. Since the degree of each vertex is at least $\frac{\varepsilon n}{4}$, it follows that

$$\mathbb{E}[Z_{ij} \mid \mathcal{F}_{i,j-1}] \leq \frac{\kappa}{100C}.$$

It follows from Azuma's inequality that

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{13}n} \sum_{j=1}^{20C} [Z_{ij} - \mathbb{E}[Z_{ij} \mid \mathcal{F}_{i,j-1}]] \geq \frac{\kappa \varepsilon_{13}}{4} n\right) \leq e^{-\kappa^2 \varepsilon_{13} n / 640C}. \tag{5.1}$$

Hence,

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{13}n} \sum_{j=1}^{20C} Z_{ij} \geq \frac{\kappa \varepsilon_{13}}{2} n\right) \leq e^{-\kappa^2 \varepsilon_{13} n / 640C}. \tag{5.2}$$

Now let B_i denote the event that the random walk starting from v_i intersects a random walk started from v_j for some $j < i$ at a point other than v_i. Let $Y_i = 1_{B_i}$. Let $\mathcal{G}_i$ denote the filtration generated by the paths of random walks started from $v_1, v_2, \dots, v_i$. Let C_i be the set of vertices visited by the first $(i-1)$ random walks except possibly v_i. Clearly, $|C_i| \le 25C\varepsilon_{13}n$. Arguing as before, the number of edges from any vertex v to C_i is at most $625C\varepsilon_{13}\log(\frac{4C_1}{25C\varepsilon_{13}})n \le \frac{\kappa\varepsilon n}{400C}$ for ε_{13} sufficiently small. By a union bound over the steps of the random walk started from v_i, it follows that $\mathbb{E}(Y_i \mid \mathcal{G}_{i-1}) \le \frac{\kappa}{4}$. Using Azuma's inequality as before, we get

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{13}n}[Y_i - \mathbb{E}[Y_i \mid \mathcal{G}_{i-1}]] \ge \frac{\kappa\varepsilon_{13}}{4}n\right) \le e^{-\kappa^2\varepsilon_{13}n/32}. \tag{5.3}$$

Hence,

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{13}n} Y_i \ge \frac{\kappa\varepsilon_{13}}{2}n\right) \le e^{-\kappa^2\varepsilon_{13}n/32}. \tag{5.4}$$

Clearly,

$$|D_1| \le \sum_{i=1}^{\varepsilon_{13}n}\sum_{j=1}^{20C} Z_{ij}$$

and

$$|D_2| \le \sum_{i=1}^{\varepsilon_{13}n} Y_i.$$

The lemma now follows from (5.2) and (5.4) and the observation at the start of the proof. $\square$

LEMMA 5.7. *Let $v_1, v_2, \dots, v_{\varepsilon_{13}n}$ be given vertices in V. From each of the vertices, we run independent continuous time random walks in $G(t)$ as described in the coupling up to time $\frac{10C}{\beta}$. Let W_1 denote the number of walks that take more than $20C$ steps in this time. Then $\mathbb{P}[W_1 \ge \kappa\varepsilon_{13}n \mid \mathcal{F}_t, t < \tau_0']$ is exponentially small in n for C sufficiently large.*

PROOF. Condition on $\{\mathcal{F}_t, t < \tau_0'\}$. Since the number of steps taken by each random walk is independent and are stochastically dominated by a Poi$(5CC_2)$ variable, it follows from a large deviation estimate and $C_2 \le 2$, that $\mathbb{P}(W_1 > \kappa\varepsilon_{13}n)$ is exponentially small in n by taking C sufficiently large. $\square$

LEMMA 5.8. *Assume the hypothesis of Lemma 5.7. Consider the coupling of the random walks with the evolving voter model as described above. Let W_2 denote the number of walkers which traverse by time $\frac{10C}{\beta}$, some edge that is rewired* *during the voter model process. Then* $\mathbb{P}(W_2 \geq 3\kappa\varepsilon_{13}n \mid \mathcal{F}_t, t < \tau_0') \leq e^{-\gamma n}$ *for* C *sufficiently large,* $\varepsilon_{13} = \varepsilon_{13}(C)$ *sufficiently small and* $\beta = \beta(C)$ *sufficiently large and for some* $\gamma > 0$ *that does not depend on* β.

PROOF. Condition on $\{\mathcal{F}_t, t < \tau_0'\}$. In the coupling construction, let ω be the number of entries in $\mathbb{RW}$ that were inspected. We consider the case $\omega \leq \frac{6Cn^2}{\beta}$ since the complement of this event has probability that is exponentially small in n^2 and can be ignored. Notice that $\mathbb{RW}$ is independent of $\mathbb{RL}$, and hence is also independent of the random walks. Let $\mathcal{H}$ denote the filtration generated by the random walk paths. Let D_i denote the set of edges traversed by the random walk started at v_i. Let $\mathcal{D}$ denote the event that there is $J \subseteq [\varepsilon_{13}n]$ with $|J| \geq (1 - 2\kappa)\varepsilon_{13}n$ such that for all $j_1, j_2 \in J$, $D_{j_1} \cap D_{j_2} = \varnothing$ and $|D_{j_1}| \leq 20C$. It follows from Lemma 5.7 and Lemma 5.6 that $\mathbb{P}(\mathcal{D}^c \mid \mathcal{F}_t, t < \tau_0')$ is exponentially small in n for C sufficiently large and $\varepsilon_{13} = \varepsilon_{13}(C)$ sufficiently small. Now let us condition on $\mathcal{H}$ and $\mathcal{D}$. Since $\mathbb{RW}$ is independent of $\mathcal{H}$, for $j \in J$, we have

$$\mathbb{P}\biggl(\text{the edge in } RW_k \in D_j \text{ for some } k \leq \frac{6Cn^2}{\beta}\biggr) \leq \frac{600C^2}{\beta}.$$

By a large deviation estimate, it follows that if β is sufficiently large so that $\beta > \frac{6000C^2}{\kappa}$ then

$$\mathbb{P}\bigl(W_2 \geq 3\kappa\varepsilon_{13}n \mid \mathcal{F}_t, t < \tau_0', \mathcal{H}, \mathcal{D}\bigr)$$

is exponentially small in n. This completes the proof of the lemma. $\square$

LEMMA 5.9. *Assume the hypothesis of Lemma 5.7. Consider the coupling of the random walks with the evolving voter model as described above. Let* W_3 *be the number of walks that take less than* $20C$ *steps up to time* $\frac{10C}{\beta}$ *but violates condition* 3 *in Definition 5.1 at time* T. *Then* $\mathbb{P}(W_3 > \kappa\varepsilon_{13}n)$ *is exponentially small in* n *for* β *sufficiently large where the exponent does not depend on* β.

PROOF. Let us fix a function $q(\cdot)$ such that $q(\beta) \ll \beta \ll q(\beta)\log q(\beta)$ as $\beta \to \infty$. Since the random walks are independent of $\mathbb{RW}$, we condition on $\mathbb{RW}$ and the following event:

$$\mathcal{G} = \biggl\{\omega < \frac{6Cn^2}{\beta}, \bigl|E^*(v)\bigr| \leq \frac{4Cnq(\beta)}{\beta}\ \forall v \in V\biggr\}.$$

By a Chernoff bound $\mathbb{P}(\mathcal{G}) \geq 1 - e^{-\gamma n}$ where γ is bounded away from 0 independent of β.

Now observe the following. For every vertex v, the chance that condition 3 in Definition 5.1 is violated while a walker takes a step from v is upper bounded by the chance that out of all the edges adjacent to v, the one to ring first is in $E^*(v)$.

On $t < \tau_0'$, the degree of v is at least $\frac{\varepsilon n}{4}$, and hence conditional on $\mathcal{G}$, this chance is upper bounded by $\frac{16Cq(\beta)}{\varepsilon\beta}$. Taking a union bound over the first $20C$ jumps made by the walk started at v_i, it follows that conditional on $\mathcal{G}$, the chance that the random walk started from v_i violates condition 3 in Definition 5.1 before making $20C$ many jumps is at most $\frac{320C^2q(\beta)}{\varepsilon\beta}$. Since the events are independent for different i conditional on $\mathcal{G}$, the lemma follows for β sufficiently large. $\square$

Now recall the coupling between the continuous time random walks on $G(t)$ and the evolving voter model defined at the start of this section. We shall need the following result for the coupling.

LEMMA 5.10. *Let $v_1, v_2, \ldots, v_{\varepsilon_{13}n}$ be fixed vertices in V. Consider the coupling between the evolving voter model starting with $G(t)$ at time t, with independent continuous time random walks on $G(t)$ starting with one walker at each v_i as described above. Let us denote the position of the random walk started at v_i at time s by $Y_i(t+s)$. Let Π denote the event that there exists a time $\sigma_0 \in \{C/\beta + \frac{1}{n^3}, C/\beta + \frac{2}{n^3}, \ldots, 3C/\beta\}$ and $J \subseteq [\varepsilon_{13}n]$ with $|J| \geq (1-8\kappa)\varepsilon_{13}n$, such that for each $i \in J$, opinion of $Y_i(t+\sigma_0)$ in $G(t)$ is same as $v_i(t + \frac{Cn^2}{\beta})$. Then for C sufficiently large, $\varepsilon_{13} = \varepsilon_{13}(C)$ sufficiently small and $\beta = \beta(C)$ sufficiently large, we have that $\mathbb{P}(\Pi \mid \mathcal{F}_t, t < \tau_0') \geq 1 - e^{-\gamma n/2}$ where γ is bounded away from 0 independent of β.*

PROOF. In this proof, the value of the constant γ may change from line to line but γ is always a positive constant bounded away from 0 independent of β. Condition on $\{\mathcal{F}_t, t < \tau_0'\}$. Let W' denote the number of walkers at time $\frac{3C}{\beta}$ of type B. Then it follows from Lemma 5.7 and Lemma 5.6 that $\mathbb{P}(W' > 2\kappa\varepsilon_{13}n) \leq e^{-\gamma n}$ for some $\gamma > 0$. Let Q denote the event that there exists $J \subseteq [\varepsilon_{13}n]$ with $|J| \geq (1-5\kappa)\varepsilon_{13}n$ such that for all $j \in J$, the opinion of $Y_j(t+T)$ in $G(t)$ is same as $v_j(t + \frac{Cn^2}{\beta})$. It now follows from Lemma 5.2, Lemma 5.8 and Lemma 5.9 that for appropriate choices of C, ε_{13} and β, $\mathbb{P}(Q^c, T < \frac{3C}{\beta} \mid \mathcal{F}_t, t < \tau_0') \leq e^{-\gamma n}$.

Now let A denote the event there exist $k \in \{1, 2, \ldots, \frac{2n^3C}{\beta}\}$ such that there are more than $\kappa\varepsilon_{13}n$ of the random walks take a step within time $[\frac{C}{\beta} + \frac{k}{n^3}, \frac{C}{\beta} + \frac{k+1}{n^3}]$. By a union bound it follows that $P(A \mid \mathcal{F}_t, t < \tau_0') \leq e^{-\gamma n}$. Since

$$\begin{aligned}
&\mathbb{P}(\Pi \mid \mathcal{F}_t, t < \tau_0') \\
&\quad \geq 1 - \mathbb{P}\Big(T \geq \frac{3C}{\beta} \Bigm| \mathcal{F}_t, t < \tau_0'\Big) \\
&\qquad - \mathbb{P}(A \mid \mathcal{F}_t, t < \tau_0') - \mathbb{P}\Big(Q^c, T < \frac{3C}{\beta} \Bigm| \mathcal{F}_t, t < \tau_0'\Big),
\end{aligned}$$

the proof of the lemma is completed using Lemma 5.3. $\square$

5.2. *Bound for large cuts.* Our aim in this section is to show that if $\{t < \tau_0\}$ and $\{t + \delta n^2 < \tau_0'\}$, then it is unlikely that the process violates the stronger threshold condition for large cuts by time $t + \delta n^2$, that is, we shall show that with high probability $\{t + \delta n^2 < \tau_2\}$.

Before proceeding with the proof, let us explain our strategy. We shall show that under the above conditions if the number of edges across any fixed cut becomes too small or too large at any time step during the evolution, then at the next step it receives a drift toward the mean, and then use a Martingale argument to show that the chance that the number of edges across that cut moves beyond the threshold is exponentially small. A union bound over all large cuts then completes the proof.

To show that the number of edges across a fixed cut is indeed mean reverting, we establish and use the fact that on $\{t + \delta n^2 < \tau_0'\}$, for most of the times in $[t, t + \delta n^2]$, the number of disagreeing edges across the cut is approximately the fraction of total number of edges across the cut that one would expect if the opinions were randomly distributed [i.e., if the number of minority opinion vertices is approximately pn, then the fraction of disagreeing edges across the cut is approximately $2p(1 - p)$]. We start by moving toward establishing this assertion using Proposition 5.5 which tells us that the opinion of the vertices become sufficiently mixed within $O(\frac{n^2}{\beta})$ steps.

5.2.1. *Fraction of disagreeing edges.* Let us fix $S, T \subseteq V$, such that $S \cap T = \varnothing$ and $S \cup T = V$ with $\varepsilon_2 n \le |S| \le |T|$. We have the following proposition.

PROPOSITION 5.11. *For the* rewire-to-random-* *dynamics, let us condition on $\{\mathcal{F}_t, t < \tau_0', N_1(t) = pn\}$. For $t' \ge t$, let $X_{ST}(t')$ denote the number of disagreeing edges at time t' with one end in S and the other in T and let $I(t')$ denote the interval $((2p(1 - p) - \varepsilon_7)N_{ST}(t'), (2p(1 - p) + \varepsilon_7)N_{ST}(t'))$. Then there exists a constant C sufficiently large, and $\beta = \beta(C)$ sufficiently large such that*

$$\mathbb{P}\biggl(X_{ST}\biggl(t + \frac{Cn^2}{\beta}\biggr) \notin I\biggl(t + \frac{Cn^2}{\beta}\biggr) \Bigm| \mathcal{F}_t, t < \tau_0'\biggr) \le e^{-\gamma n}$$

for some $\gamma > 0$ that does not depend on β.

We shall need the following lemma in order to prove Proposition 5.11.

LEMMA 5.12. *Let $\mathcal{E}_{S,T}(t)$ denote the set of edges that have one endpoint in S and another endpoint T at time t. On $\{t < \tau_2'\}$, and hence on $\{t < \tau_0'\}$, we have $|\mathcal{E}_{ST}(t)| \ge \frac{\varepsilon_2 n^2}{5}$ provided $1000\varepsilon_3 < \varepsilon_2$. Let $\kappa > 0$ be fixed. Let $e_1, e_2, \dots, e_{\varepsilon_{12} n}$ be uniformly chosen edges from $\mathcal{E}_{ST}(t)$. For ε_{12} sufficiently small, with exponentially high probability there exists at least $(1 - \kappa)\varepsilon_{12} n$ many edges among the sample that are vertex disjoint at time t.*

PROOF. For $i = 1, 2, \ldots, \varepsilon_{12}n$, let A_i denote the event that e_i is not vertex disjoint with $e_1, e_2, \ldots, e_{i-1}$. Let $Z_i = 1_{A_i}$. Let $\mathcal{G}_{i-1}$ denote the filtration generated by $e_1, e_2, \ldots, e_{i-1}$. Then it is clear from the assumption on the graph that

$$\mathbb{E}(Z_i \mid \mathcal{G}_{i-1}) \le \frac{C_2 n}{|\mathcal{E}_{ST}(t)|} 2\varepsilon_{12} n \le \frac{12 C_2 \varepsilon_{12}}{\varepsilon_2}.$$

Also notice that $|Z_i - \mathbb{E}(Z_i \mid \mathcal{G}_{i-1})| \le 1$, and hence Azuma's inequality yields

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{12}n} [Z_i - \mathbb{E}[Z \mid \mathcal{G}_{i-1}]] \ge \frac{\kappa \varepsilon_{12}}{2} n\right) \le e^{-\kappa^2 \varepsilon_{12} n/8}. \tag{5.5}$$

It follows that

$$\mathbb{P}\left(\sum_{i=1}^{\varepsilon_{12}n} Z_i \ge \frac{12 C_2 \varepsilon_{12}^2 n}{\varepsilon_2} + \frac{\kappa \varepsilon_{12}}{2} n\right) \le e^{-\kappa^2 \varepsilon_{12} n/8}. \tag{5.6}$$

By choosing ε_{12} sufficiently small such that $\frac{12C_2\varepsilon_{12}}{\varepsilon_2} \le \frac{\kappa}{2}$ completes the proof. $\square$

Now we continue with the proof of Proposition 5.11.

PROOF OF PROPOSITION 5.11. Let us condition on $\{\mathcal{F}_t, t < \tau_0', N_1(t) = pn\}$. Let $N_{ST}(t)$ be the number of edges in $G(t)$ with one endpoint in S and another endpoint in T. Let $\{e_1, e_2, \ldots, e_{N_{ST}(t)}\}$ denote the set of those edges. Let X_i be the indicator that endpoints of e_i in $G(t)$ are disagreeing in $G(t + \frac{Cn^2}{\beta})$. Observe that in $\frac{Cn^2}{\beta}$ steps, at most $\frac{Cn^2}{\beta}$ edges can be rewired. Hence, we have $|N_{ST}(t + \frac{Cn^2}{\beta}) - N_{ST}(t)| \le \frac{Cn^2}{\beta}$ and also

$$\left| X_{ST}\left(t + \frac{Cn^2}{\beta}\right) - \sum_{i=1}^{N_{ST}(t)} X_i \right| \le \frac{Cn^2}{\beta}.$$

It follows using $N_{ST}(t) \ge \frac{\varepsilon_2 n^2}{5}$ that for β sufficiently large, the difference between $\frac{X_{ST}(t+\frac{Cn^2}{\beta})}{N_{ST}(t+\frac{Cn^2}{\beta})}$ and $\frac{\sum X_i}{N_{ST}(t)}$ can be made smaller than $\varepsilon_7/2$ and hence it suffices to prove that

$$\mathbb{P}\left(\frac{1}{N_{ST}(t)} \sum_{i=1}^{N_{ST}(t)} X_i \notin \left(2p(1-p) - \frac{\varepsilon_7}{2},\right.\right.$$
$$\left.\left. 2p(1-p) + \frac{\varepsilon_7}{2}\right) \,\middle|\, \mathcal{F}_t, t < \tau_0'\right) \le e^{-\gamma n}.$$

Let J be a set of size $\varepsilon_{12}n$ where each element is an independent uniform sample from $[N_{ST}(t)]$. Clearly, by Hoeffding's inequality,

$$\mathbb{P}\left(\left|\frac{1}{\varepsilon_{12}n}\sum_{j\in J}X_j-\frac{1}{N_{ST}(t)}\sum_{i=1}^{N_{ST}(t)}X_i\right|\geq\frac{\varepsilon_7}{4}\right)\leq e^{-\frac{\varepsilon_7^2\varepsilon_{12}n}{32}}.$$

So it suffices for us to prove that with probability at least $1-e^{-2\gamma n}$,

$$\frac{1}{\varepsilon_{12}n}\sum_{j\in J}X_j\in\big(2p(1-p)-\varepsilon_7/4,2p(1-p)+\varepsilon_7/4\big).$$

Choose ε_{12} sufficiently small, and set $\varepsilon_{13}=2\varepsilon_{12}(1-\kappa)$ so that the conclusions of Lemma 5.12 and Lemma 5.10 are satisfied. Let $\mathcal{H}_1$ denote the event that there is a subset $J^*\subseteq J$ with $|J^*|=(1-\kappa)\varepsilon_{12}n$ such that endpoints of e_j are disjoint for all $j\in J^*$. It follows from Lemma 5.12 that $\mathbb{P}(\mathcal{H}_1\mid\mathcal{F}_t,t<\tau_0')\geq 1-e^{-100\gamma n}$. Condition on $\mathcal{H}_1$ and J^*. Let $v_1,v_2,\dots,v_{\varepsilon_{13}n}$ be endpoints of edges of J^*. By choosing $100\kappa<\varepsilon\varepsilon_7$ it follows that it suffices to prove

$$\begin{aligned}\mathbb{P}\Bigg[\frac{2}{\varepsilon_{13}n}\sum_{j\in J^*}X_j\notin\big(2p(1-p)-\varepsilon_7/8,2p(1-p)+\varepsilon_7/8\big)\mid\\ \mathcal{F}_t,t<\tau_0',N_1(t)=pn,\mathcal{H}_1,J\Bigg]\leq e^{-10\gamma n}.\end{aligned}\tag{5.7}$$

Consider the coupling described in Section 5.1. Fix $j\in J^*$, let v_{j_1} and v_{j_2} be endpoints of e_j in $G(t)$. For $\tilde{\sigma}\in\Sigma=\{C/\beta+\frac{1}{n^3},C/\beta+\frac{2}{n^3},\dots,3C/\beta\}$, let $U_{j,\tilde{\sigma}}$ denote the indicator that the position of the coupled random walks started from v_{j_1} and v_{j_2} at time $\tilde{\sigma}$ have different opinions in $G(t)$. Clearly, for a fixed $\tilde{\sigma}$, for all $j\in J^*$, $U_{j,\tilde{\sigma}}$ are conditionally independent. Also, it follows from Proposition 5.5 that $\mathbb{E}(U_{j,\tilde{\sigma}}\mid\mathcal{F}_t,t<\tau_0',N_1(t)=pn)\in[2p(1-p)-\varepsilon_7/32,2p(1-p)+\varepsilon_7/32]$. A standard Hoeffding bound now shows that conditional on $\mathcal{G}=\{\mathcal{F}_t,t<\tau_0',N_1(t)=pn,\mathcal{H}_1,J^*\}$, with probability at least $1-e^{-20\gamma n}$,

$$\frac{2}{\varepsilon_{13}n}\sum_{j\in J^*}U_{j,\tilde{\sigma}}\in\big(2p(1-p)-\varepsilon_7/16,2p(1-p)+\varepsilon_7/16\big).\tag{5.8}$$

By taking a union bound over all possible values of $\tilde{\sigma}$, it follows that above holds for all $\tilde{\sigma}$ in Σ with conditional probability at least $1-e^{-15\gamma n}$.

By observing that by Lemma 5.10, we have, conditional on $\mathcal{G}$, there exists $\tilde{\sigma}\in\Sigma$

$$\frac{2}{\varepsilon_{13}n}\sum_{j\in J^*}|U_{j,\tilde{\sigma}}-X_j|\leq 12\kappa\leq\varepsilon_7/32$$

with probability at least $1-e^{-15\gamma n}$. This and the previous observation imply (5.7) and the proof of the proposition is complete. □

We shall also need to consider the fraction of disagreeing edges with both endpoints in S. The following proposition follows along the same lines as Proposition 5.11 and we shall omit the proof.

PROPOSITION 5.13. *Fix $S \subseteq V$ with $|S| \geq \varepsilon_2 n$. For the* rewire-to-random-* *dynamics, let us condition on $\{\mathcal{F}_t, t < \tau_0', N_1(t) = pn\}$. For $t' \geq t$, let $X_{SS}(t')$ denote the number of disagreeing edges at time t' and let $I'(t')$ denote the interval $((2p(1-p) - \varepsilon_7)N_{SS}(t'), (2p(1-p) + \varepsilon_7)N_{SS}(t'))$. Then there exists a constant C sufficiently large, and $\beta = \beta(C)$ sufficiently large such that*

$$\mathbb{P}\left(X_{SS}\left(t + \frac{Cn^2}{\beta}\right) \notin I'\left(t + \frac{Cn^2}{\beta}\right) \,\Big|\, \mathcal{F}_t, t < \tau_0'\right) \leq e^{-\gamma n}$$

for some $\gamma > 0$ that does not depend on β.

Fix S and T as in Proposition 5.11. We shall need to number of times in $[t + 1, t + \delta n^2]$ for which among the edges across (S, T), within S and within T each has the "correct" fraction of disagreeing edges. We shall show that the probability that this fraction of times is bigger than a small constant is exponentially small. There is a subtle point here. We will ultimately want to take a union bound over all large cuts (S, T) which are exponentially many in number. Hence, we need to make sure that the exponent one gets in the exponentially small probability described above is sufficiently large to offset this large union bound. To this end, we shall use the fact that the time taken for the opinions to mix is $O(\frac{n^2}{\beta}) \ll \delta n^2$ which can be ensured by taking β sufficiently large. We now move toward making the above discussion quantitative.

Condition on $\{\mathcal{F}_t, t < \tau_0, N_1(t) = pn\}$. Let $D^*(t')$ denote the event that $N_1(t'') \in ((p - \varepsilon_7)n, (p + \varepsilon_7)n)$ for all $t'' \in [t + 1, t']$. Fix S and T as in Proposition 5.11. For $t' \in [t + 1, t + \delta n^2]$, let us define events $\mathcal{A}_{ST}^{t'}$, $\mathcal{A}_{SS}^{t'}$ and $\mathcal{A}_{TT}^{t'}$ as follows:

$$\mathcal{A}_{ST}^{t'} = \{X_{ST}(t') \in (2p(1-p) - 2\varepsilon_7, 2p(1-p) + 2\varepsilon_7)N_{ST}(t')\},$$

$$\mathcal{A}_{SS}^{t'} = \{X_{SS}(t') \in (2p(1-p) - 2\varepsilon_7, 2p(1-p) + 2\varepsilon_7)N_{SS}(t')\},$$

$$\mathcal{A}_{TT}^{t'} = \{X_{TT}(t') \in (2p(1-p) - 2\varepsilon_7, 2p(1-p) + 2\varepsilon_7)N_{TT}(t')\}.$$

Finally, let us define

$$\mathcal{A}^{t'} = \mathcal{A}_{SS}^{t'} \cap \mathcal{A}_{ST}^{t'} \cap \mathcal{A}_{TT}^{t'}.$$

Let $Z_{t'}$ be the indicator of $\overline{\mathcal{A}^{t'}}$, that is, the complement of $\mathcal{A}^{t'}$. We have the following lemma controlling the fraction of times t' for which $\mathcal{A}^{t'}$ hold.

LEMMA 5.14. *Set $\mathcal{G} = \{\mathcal{F}_t, t < \tau_0, N_1(t) = pn\}$. Then for each $\varepsilon_{15} > 0$ we have*

$$\mathbb{P}\left(\frac{1}{\delta n^2}\sum_{t'=t+1}^{t+\delta n^2} Z_{t'} \geq \varepsilon_{15}, t+\delta n^2 < \tau_0', D^*(t+\delta n^2) \mid \mathcal{G}\right) \leq e^{-h(\beta)n},$$

where $h(\beta)$ can be made arbitrarily large by taking β sufficiently large.

PROOF. For $i = 1, 2, \ldots, \frac{Cn^2}{\beta}$, and for $k = 1, 2, \ldots, \frac{\delta\beta}{C} - 1$ let $t_{i,k} = t + i + k\frac{Cn^2}{\beta}$. Let

$$W_i = \sum_{k=1}^{\frac{\delta\beta}{C}-1} Z_{t_{i,k}}.$$

From Proposition 5.11 and Proposition 5.13, it follows that for a fixed i, $\mathbb{P}[Z_{t_{i,k+1}} = 1 \mid \mathcal{F}_{t_{i,k}}, t_{i,k} < \tau_0', D^*(t_{i,k})] \leq e^{-\gamma n/5}$. Observe that even though $Z_{t_i,k}$ are not independent (as k varies), the above bound implies that they are jointly stochastically dominated by an independent collection of indicators and in particular one can bound the exponential moment of W_i, in exactly the same manner in the standard proof of Chernoff's inequality, and hence the standard Chernoff bound applies in this case. Hence, we have

$$\begin{aligned}
&\mathbb{P}\left[\frac{C}{\delta\beta} W_i \geq \varepsilon_{15}/2, t+\delta n^2 \leq \tau_0', D^*(t+\delta n^2) \mid \mathcal{G}\right] \\
&\quad \leq \exp\left(-\varepsilon_{15}\delta\beta \log\left(\frac{\varepsilon_{15} e^{-\gamma n/5}}{2}\right)\Big/ 4C\right) \qquad (5.9)\\
&\quad \leq \left(\frac{2}{\varepsilon_{15}}\right)^{1/4C} \exp(-\varepsilon_{15}\delta\beta\eta n/20C).
\end{aligned}$$

Taking a union bound over all i and choosing β sufficiently large so that $\frac{C}{\delta\beta} \leq \varepsilon_{15}/2$, it follows that

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{\delta n^2}\sum_{t'=t+1}^{t+\delta n^2} Z_{t'} \geq \varepsilon_{15}, t+\delta n^2 < \tau_0', D^*(t+\delta n^2) \mid \mathcal{F}_t\right) \\
&\quad \leq \frac{Cn^2}{\beta}\left(\frac{2}{\varepsilon_{15}}\right)^{1/4C} \exp(-\varepsilon_{15}\delta\beta\gamma n/20C).
\end{aligned}$$

Taking β sufficiently large completes the proof of the lemma. $\square$

5.2.2. *Obtaining a bound for τ_2.* Now we move toward the main result in this subsection, that is, showing that the evolving voter model process is unlikely to reach τ_2 before time $t + \delta n^2$. We start with a fixed cut (S, T) as above with $\varepsilon_2 n \le |S| \le |T|$. Recall that

$$K_{ST}(t') = \left(\frac{N_{SS}(t') - \frac{1}{4}|S|^2}{N(t')}\right)^2 + \left(\frac{N_{TT}(t') - \frac{1}{4}|T|^2}{N(t')}\right)^2.$$

We have the following proposition.

PROPOSITION 5.15. *Set $\mathcal{G} = \{\mathcal{F}_t, t < \tau_0, N_1(t) = pn\}$. Let S, T be as above. Then we have*

$$\mathbb{P}\bigl(K_{ST}(t + \delta n^2) > \varepsilon_3^2, t + \delta n^2 < \tau_0', D^*(t + \delta n^2) \mid \mathcal{G}\bigr) \le \exp\bigl(-h(\beta)n\bigr),$$

where $h(\beta)$ can be made arbitrarily large by taking β sufficiently large.

PROOF. For $s \in [t, t + \delta n^2 - 1]$, let $\mathcal{F}_s$ denote the filtration generated by the process up to time s. Conditioned on $\mathcal{F}_s$ the transition rule for the evolution of $(N_{SS}(s), N_{TT}(s))$ is given by the following:

$$\begin{aligned}
&\bigl(N_{SS}(s+1), N_{TT}(s+1)\bigr) \\
&\quad = \begin{cases}
\bigl(N_{SS}(s) + 1, N_{TT}(s)\bigr) & \text{w.p.} \quad \dfrac{X_{ST}(s)}{N(s)} \dfrac{1}{2} \dfrac{|S| - 1}{n - 1}\left(1 - \dfrac{\beta}{n}\right), \\
\bigl(N_{SS}(s), N_{TT}(s) + 1\bigr) & \text{w.p.} \quad \dfrac{X_{ST}(s)}{N(s)} \dfrac{1}{2} \dfrac{|T| - 1}{n - 1}\left(1 - \dfrac{\beta}{n}\right), \\
\bigl(N_{SS}(s) - 1, N_{TT}(s)\bigr) & \text{w.p.} \quad \dfrac{X_{SS}(s)}{N(s)} \dfrac{|T|}{n - 1}\left(1 - \dfrac{\beta}{n}\right), \\
\bigl(N_{SS}(s), N_{TT}(s) - 1\bigr) & \text{w.p.} \quad \dfrac{X_{TT}(s)}{N(s)} \dfrac{|S|}{n - 1}\left(1 - \dfrac{\beta}{n}\right), \\
\bigl(N_{SS}(s), N_{TT}(s)\bigr) & \text{otherwise.}
\end{cases}
\end{aligned}$$

For this proof, let us write $\Delta = 2p(1-p)(1 - \frac{\beta}{n})$ and $\varepsilon_8 = \varepsilon_7 / 2p(1-p)$. Recall that on $Z_s = 0$, we have that the ratios $\frac{X_{SS}(s)}{N_{SS}(s)}, \frac{X_{ST}(s)}{N_{ST}(s)}, \frac{X_{TT}(s)}{N_{TT}(s)}$ each belong to the interval $(2p(1-p) - \varepsilon_7, 2p(1-p) + \varepsilon_7)$. Hence, for n sufficiently large we have on $\{Z_s = 0\}$

$$\begin{aligned}
&\bigl(N_{SS}(s+1), N_{TT}(s+1)\bigr) \\
&\quad = \bigl(N_{SS}(s) + 1, N_{TT}(s)\bigr) \qquad \text{w.p.} \quad \in \Delta \frac{N_{ST}(s)}{N(s)} \frac{1}{2} \frac{|S|}{n}(1 \pm 3\varepsilon_8)
\end{aligned}$$

and the other transition probabilities above can be expressed similarly.

Doing a change of variable

$$W_S(s) = \frac{N_{SS}(s) - |S|^2/4}{N}$$

and

$$W_T(s) = \frac{N_{TT}(s) - |T|^2/4}{N}$$

it follows that on $\{Z_s = 0\}$, we have

$$\begin{aligned}
&\bigl(W_S(s+1), W_T(s+1)\bigr) \\
&\quad = \begin{cases}
\bigl(W_S(s) + 1/N, W_T(s)\bigr) \\
\qquad \text{w.p.} \quad \in \Delta\biggl(1 - W_S(s) - W_T(s) - \dfrac{|S|^2 + |T|^2}{4N}\biggr)\dfrac{|S|}{2n} \\
\qquad \pm 3\varepsilon_7, \\
\bigl(W_S(s), W_T(s) + 1/N\bigr) \\
\qquad \text{w.p.} \quad \in \Delta\biggl(1 - W_S(s) - W_T(s) - \dfrac{|S|^2 + |T|^2}{4N}\biggr)\dfrac{|T|}{2n} \\
\qquad \pm 3\varepsilon_7, \\
\bigl(W_S(s) - 1/N, W_T(s)\bigr) \\
\qquad \text{w.p.} \quad \in \Delta\biggl(W_S(s) + \dfrac{|S|^2}{4N}\biggr)\dfrac{|T|}{n} \pm 3\varepsilon_7, \\
\bigl(W_S(s), W_T(s) - 1/N\bigr) \\
\qquad \text{w.p.} \quad \in \Delta\biggl(W_T(s) + \dfrac{|T|^2}{4N}\biggr)\dfrac{|S|}{n} \pm 3\varepsilon_7.
\end{cases}
\end{aligned}$$

It follows that on $\{Z_s = 0\}$,

$$\begin{aligned}
&\mathbb{E}\bigl[W_S(s+1)^2 \mid \mathcal{F}_s\bigr] \\
&\quad = W_S(s)^2 + \frac{2W_S(s)}{N}\biggl(\Delta\frac{|S|}{2n}\bigl(1 - W_S(s) - W_T(s)\bigr) \\
&\quad - \frac{2W_S(s)}{N}\biggl(\frac{|S|^2 + |T|^2}{4N}\biggr) - \Delta\frac{|T|}{n}\biggl(W_S(s) + \frac{|S|^2}{4N}\biggr) \pm 6\varepsilon_7\biggr) \\
&\qquad + O\biggl(\frac{1}{N^2}\biggr) \\
&\quad = W_S(s)^2 + \frac{2W_S(s)\Delta}{N}\biggl[\frac{|S|}{2n} - \frac{|S|n}{8N} - W_S(s)\biggl(\frac{|S| + 2|T|}{2n}\biggr) - W_T(s)\frac{|S|}{2n}\biggr] \\
&\qquad\quad + o\biggl(\frac{1}{n^2}\biggr) \pm \frac{25\varepsilon_7}{n^2}.
\end{aligned}$$

Doing a similar calculation for $W_T(s+1)^2$ we get, on $\{Z_s=0\}$,

$$
\begin{aligned}
&\mathbb{E}\big[W_S(s+1)^2+W_T(s+1)^2 \mid \mathcal{F}_s\big]\\
&\quad \le W_S(s)^2+W_T(s)^2\\
&\qquad + \frac{2W_S(s)\Delta}{N}\left[\frac{|S|}{2n}-\frac{|S|n}{8N}-W_S(s)\left(\frac{|S|+2|T|}{2n}\right)-W_T(s)\frac{|S|}{2n}\right]\\
&\qquad + \frac{2W_T(s)\Delta}{N}\left[\frac{|T|}{2n}-\frac{|T|n}{8N}-W_T(s)\left(\frac{2|S|+|T|}{2n}\right)-W_T(s)\frac{|T|}{2n}\right]\\
&\qquad + \frac{50\varepsilon_7}{n^2}+o\left(\frac{1}{n^2}\right)\\
&\quad \le W_S(s)^2+W_T(s)^2\\
&\qquad - \frac{4\Delta}{n^2}\left[\frac{1}{4}\big(W_S(s)^2+W_T(s)^2\big)+\frac{1}{4}\big(W_s(s)+W_T(s)\big)^2\right]\\
&\qquad - \frac{4\Delta}{n^2}\left[\frac{|T|}{2n}W_S(s)^2+\frac{|S|}{2n}W_T(s)^2\right]+\frac{64\varepsilon_7}{n^2}\\
&\quad \le W_S(s)^2+W_T(s)^2-\frac{\Delta}{n^2}\big[\big(W_S(s)^2+W_T(s)^2\big)-64\varepsilon_8\big].
\end{aligned}
$$

Recall that

$$K_{ST}(t)=\left(\frac{N_{SS}(t)-\frac{1}{4}|S|^2}{N(t)}\right)^2+\left(\frac{N_{TT}(t)-\frac{1}{4}|T|^2}{N(t)}\right)^2.$$

Hence, we have from above

$$\mathbb{E}\big[K_{ST}(s+1)-K_{ST}(s) \mid \mathcal{F}_s, Z_s=0\big] \le -\frac{\Delta}{n^2}\big(K_{ST}(s)-64\varepsilon_8\big).$$

In particular, on $\{Z_s=0\}\cap\{K_{ST}(s)\ge \varepsilon_3^2/2\}$, we have

$$\mathbb{E}\big[K_{ST}(s+1)-K_{ST}(s) \mid \mathcal{F}_s\big] \le -\frac{\Delta}{4n^2}\varepsilon_3^2$$

by choosing $\varepsilon_3^2 \ge 256\varepsilon_8$. Let $\mathcal{C}$ denote the event

$$\mathcal{C}=\Big\{\min_{s\in[t,t+\delta n^2]} K_{ST}(s)\le \varepsilon_3^2/2\Big\}.$$

Let $\mathcal{D}$ denote the event

$$\mathcal{D}=\left\{\sum_{s=t}^{t+\delta n^2-1}\mathbb{E}\big[K_{ST}(s+1)-K_{ST}(s) \mid \mathcal{F}_s\big] > \delta\big(16\varepsilon_{15}-(1-\varepsilon_{15})\Delta\varepsilon_3^2/4\big)\right\}.$$

Observe that on $\{D^*(t+\delta n^2), t+\delta n^2<\tau_0'\}$, we have that $\mathcal{C}^c\cap\mathcal{D}$ implies that the number of $s\in[t,t+\delta n^2-1]$ such that $Z_s=1$, is at least $\varepsilon_{15}\delta n^2$. It follows now from Lemma 5.14 that

$$\mathbb{P}[\mathcal{C}^c, \mathcal{D}, D^*(t+\delta n^2), t+\delta n^2 < \tau_0' \mid \mathcal{G}] \le \exp(-h(\beta)n),$$

where $h(\beta)$ can be made arbitrarily large by choosing β sufficiently large.

Now an application of Azuma–Hoeffding inequality gives

$$\begin{aligned}\mathbb{P}\Bigg[\sum_{s=t}^{t+\delta n^2-1} & K_{ST}(s+1) - K_{ST}(s) \\ & - \mathbb{E}(K_{ST}(s+1) - K_{ST}(s) \mid \mathcal{F}_s) \ge \frac{\delta\varepsilon_3^2}{10} \Bigm| \mathcal{G}\Bigg] \\ & \le \exp\left(-\frac{\delta\varepsilon_3^4 n^2}{204{,}800}\right).\end{aligned} \tag{5.10}$$

It follows that if ε_{15} is chosen sufficiently small so that $\varepsilon\varepsilon_3^2 > \frac{64\varepsilon_{15}}{1-\varepsilon_{15}}$, then we have

$$\begin{aligned}&\mathbb{P}[K_{ST}(t+\delta n^2) > K_{ST}(t), \mathcal{C}^c, D^*(t+\delta n^2)t+\delta n^2 < \tau_0' \mid \mathcal{G}] \\ &\quad \le \exp(-h(\beta)n).\end{aligned}$$

Now notice that $K_{ST}(s+1) - K_{ST}(s) \le \frac{16}{n^2}$. Observe further that by choosing δ sufficiently small ($16\delta < \varepsilon_3^2$), it follows that on $\mathcal{C}$, $K_{ST}(t+\delta n^2) < \varepsilon_3^2$. Hence,

$$\begin{aligned}&\mathbb{P}[K_{ST}(t+\delta n^2) > \varepsilon_3^2, D^*(t+\delta n^2), t+\delta n^2 < \tau_0' \mid \mathcal{G}] \\ &\quad \le \exp(-h(\beta)n).\end{aligned}$$

This completes the proof of the proposition. □

Now we are ready to prove the main result of this subsection.

THEOREM 5.16. *We have for all* $t \ge \delta n^2$,

$$\mathbb{P}[t+\delta n^2 \ge \tau_2, t+\delta n^2 < \tau_*, t < \tau_0] \le \frac{1}{n^{14}}.$$

PROOF. Let $\tilde{D}(t+\delta n^2)$ be defined as follows:

$$\tilde{D}(t+\delta n^2) = \{\forall t' \in [t+1, t+\delta n^2] : |N_1(t') - N_1(t)| \le \varepsilon_7 n\}.$$

Recall that

$$L(t) = \max_{S,T:\min(|S|,|T|)\ge\varepsilon_2 n} K_{ST}(t).$$

By taking a union bound over all cuts S, T such that $\varepsilon_2 n \le |S| \le |T|$ and using Proposition 5.15 we get that

$$\mathbb{P}\big[L\big(t+\delta n^2\big) \ge \varepsilon_3^2, \tilde{D}\big(t+\delta n^2\big), t+\delta n^2 < \tau_0' \mid \mathcal{F}_t, t < \tau_0\big] \le 2^n \exp\big(-h(\beta)n\big) \le \frac{1}{n^{18}}$$

by taking β sufficiently large. It follows by a random walk estimate that $\mathbb{P}[\tilde{D}(t + \delta n^2) \mid \mathcal{F}_t, t < \tau_0]$ is exponentially close to 1, and hence we have

$$\mathbb{P}\big[L\big(t + \delta n^2\big) \ge \varepsilon_3^2, t + \delta n^2 < \tau_0' \mid t < \tau_0, \mathcal{F}_t\big] \le \frac{2}{n^{18}}.$$

By Theorem 4.9, we know that $\mathbb{P}[t + \delta n^2 \ge \tau_0', t + \delta n^2 < \tau_* \mid \mathcal{F}_t, t < \tau_0] \le \frac{1}{n^{17}}$. Since $\{L(t + \delta n^2) \ge \varepsilon_3^2, t + \delta n^2 < \tau\}$ is contained in the union of $\{t + \delta n^2 \ge \tau_0', t + \delta n^2 < \tau_*\}$ and $\{L(t + \delta n^2) \ge \varepsilon_3^2, t + \delta n^2 < \tau_0'\}$, it follows that

$$\mathbb{P}\big[L\big(t + \delta n^2\big) \ge \varepsilon_3^2, t + \delta n^2 < \tau_* \mid \mathcal{F}_t, t < \tau_0\big] \le \frac{2}{n^{17}}.$$

Now to infer $\tau_2 > t + \delta n^2$, we need to show that $L(t + \delta n^2 - s) \le \varepsilon_3^2$ for each $s \in [0, \delta n^2 - 1]$. As $\{t < \tau_0\} \subseteq \{t - s < \tau_0\}$ and $\{t + \delta n^2 < \tau_*\} \subseteq \{t + \delta n^2 - s < \tau_*\}$ for each $s \ge 0$, by taking a union over $s \in [0, \delta n^2 - 1]$ we get that

$$\mathbb{P}\big[\tau_2 \le t + \delta n^2, t + \delta n^2 < \tau_*, t < \tau_0\big] \le \frac{1}{n^{14}}.$$

This completes the proof of the theorem. □

5.3. *Edge multiplicity.* In this subsection, we consider the strong stopping time associated with edge multiplicities, that is, τ_3. We shall prove that if $t < \tau_0$, then it is unlikely for the process to reach τ_3 by time $t + \delta n^2$.

THEOREM 5.17. *Let $t < n^4$. Then $\mathbb{P}[t \ge \tau_3, t < \tau_0'] \le \frac{1}{n^4}$.*

To prove Theorem 5.17, we shall argue that if a bond contains too many edges then it becomes likely for it to lose edges as well. To establish this, we shall need to show that each bond spends a positive fraction of the time in $[t, t + \delta n^2]$ being disagreeing. We start with the following lemma.

LEMMA 5.18. *Let u, v be two fixed vertices in V. Let $X_{uv}(t')$ be the indicator that u and v are disagreeing in $G(t')$. Then we have $\mathbb{E}(X_{uv}(t + \frac{Cn^2}{\beta}) \mid \mathcal{F}_t, t < \tau_0') \ge \frac{\varepsilon}{2}$.*

PROOF. To prove this, consider the coupling of the evolving voter model with independent continuous time random walk started from u and v as described in Section 5.1, and recall that $\mathbb{T}$ corresponds, in continuous time, to the time $t + \frac{Cn^2}{\beta}$.

Let Y_{uv} denote the indicator that the positions of the random walks started from u and v after time $\mathbb{T}$ are disagreeing in $G(t)$. We claim that $\mathbb{P}(X_{uv}(t+\frac{Cn^2}{\beta}) \neq Y_{uv}) \leq \varepsilon/4$. Indeed as in Lemma 5.2, it suffices to prove that the chance that any of the following three events occur is at most $\frac{\varepsilon}{4}$: (a) the random walks started from u and v intersect up to time $\mathbb{T}$, (b) either of the random walks traverse an edge that was rewired and (c) either of the walks violate condition 3 in Definition 5.1.

Notice that arguing as in Lemma 5.6, the chance that the random walks intersect up to time $\frac{3C}{\beta}$ is $o(1)$ as $n \to \infty$. Also notice that, following the argument in Lemma 5.8 the chance that either of the walk traverses any edge that was rewired by time $\frac{3C}{\beta}$ can be made less than $\frac{\varepsilon}{100}$ by choosing β sufficiently large. Arguing as in the proof of Lemma 5.9, it follows that for β sufficiently large the chance that either of the walks violate condition 3 in Definition 5.1 is also less than $\frac{\varepsilon}{100}$. The claim follows by noting that with high probability $\mathbb{T} \leq \frac{3C}{\beta}$ by Lemma 5.3.

Also let Y^*_{uv} be the indicator that the position of random walks started from u and v after time $\frac{2C-0.01\varepsilon}{\beta}$ are disagreeing in $G(t)$. Observe that using Lemma 5.3 it follows that the chance that either of the random walks took a step during time $[\frac{2C-0.01\varepsilon}{\beta}, \mathbb{T}]$ is bounded by $\frac{\varepsilon}{100}$, and hence $\mathbb{P}(Y_{uv} \neq Y^*_{uv}) \leq \varepsilon/4$. The lemma now follows by noticing that Proposition 5.5 implies that for C sufficiently large $\mathbb{E}(Y^*_{uv}) \geq \varepsilon$. $\square$

LEMMA 5.19. *Let $u, v \in V$ be two vertices in V. Fix $t > \varepsilon_{16} n^2 \log n$. For $\varepsilon_{16} n^2 \log n < t' < t$, let $A_{t'}$ denote the event that there exists $\mathcal{T} \in \{1, 2, \dots, t-t'\}$ such that*

$$\#\{s \in \{\mathcal{T}, \mathcal{T}+1, \dots, \mathcal{T}+t'\} : u(s) \neq v(s)\} \leq t'\varepsilon/4.$$

Then we have

$$\mathbb{P}\Big[\bigcup_{t'} A_{t'}, t < n^4 \wedge \tau'_0\Big] \leq \frac{1}{n^{r(\beta)}},$$

where $r(\beta)$ can be made arbitrarily large by taking β sufficiently large.

PROOF. Fix $\varepsilon_{16} n^2 \log n < t' < t \wedge n^4$ and $\mathcal{T} \in \{1, 2, \dots, t-t'\}$. For $t'' \in \{\mathcal{T}, \mathcal{T}+1, \dots, \mathcal{T}+t'\}$ it follows from Lemma 5.18 that $\mathbb{P}[u(t''+Cn^2/\beta) \neq v(t''+Cn^2/\beta) \mid \mathcal{F}_{t''}, t'' < \tau'_0] \geq \varepsilon/2$ for β sufficiently large. It follows using a Chernoff's bound that for each $i = 1, 2, \dots, Cn^2/\beta$,

$$\begin{aligned}
&\mathbb{P}\Big[\#\{t'' \in \{\mathcal{T}+i+kCn^2/\beta : k \in [\beta t'/n^2 C]\} : u(t'') \neq v(t'')\} \\
&\qquad \leq \frac{\beta t'\varepsilon}{4Cn^2}, t < \tau'_0\Big] \qquad (5.11)\\
&\qquad \leq \exp\Big(-\frac{\beta t'\varepsilon}{12Cn^2}\Big).
\end{aligned}$$

For all $t' > \varepsilon_{16} n^2 \log n$, we have the right-hand side of the above inequality is at most $(\frac{1}{n})^{\beta\varepsilon_{16}\varepsilon/12C}$. Let $A^*_{t',\mathcal{T},i}$ denote the event

$$A^*_{t',\mathcal{T},i} = \Big\{\#\{t'' \in \{\mathcal{T} + i + kCn^2/\beta : k \in [\beta t'/n^2 C]\} : u(t'') \neq v(t'')\} \\ \leq \frac{\beta t' \varepsilon}{4Cn^2}\Big\}.$$

Observe that on $\bigcap_i A^*_{t',\mathcal{T},i}$, the number of times $s \in \{\mathcal{T}, \mathcal{T}+1, \dots, \mathcal{T}+t'\}$ is at least $\frac{\varepsilon t'}{4}$, and hence it follows by taking a union bound over all $i \in [Cn^2/\beta]$ and all $\mathcal{T} \in \{1, 2, \dots, t - t'\}$ and using (5.11) that

$$\mathbb{P}[A_{t'}, t < n^4 \wedge \tau_0'] \leq \frac{1}{n^{r'(\beta)}},$$

where $r'(\beta)$ can be made sufficiently large by choosing β to be sufficiently large. The lemma now follows by taking union bound over $t' \in (\varepsilon_{16} n^2 \log n, t \wedge n^4)$. □

Now we define the following family of random walks which we couple with the *rewire-to-random-** dynamics as follows, $X^s(\cdot)$ indexed by $s \in \{1, 2, \dots, t\}$ with each starting from $K > 0$ [i.e., $X^s(0) = K\ \forall s$] with transition probabilities as described below:

$$X^s(h+1) = \begin{cases} X^s(h) + 1 & \text{w.p. } \dfrac{9C_2}{n^2}, \\ X^s(h) - 1 & \text{w.p. } \dfrac{K}{n^2} \text{ if } u(s+h) \neq v(s+h), \\ X^s(h) & \text{otherwise.} \end{cases}$$

The following lemma is immediate by comparing one step transition probabilities of $M_{uv}(t)$ and $X^s(t-s)$.

LEMMA 5.20. *Let* $M^*_{uv}(t) = \max_{t' \in [1,t]} M_{uv}(t')$ *and* $X^*(t) = \max_{s,h: s+h \leq t} X^s(h)$. *Then we have, on* $\{t < \tau_0'\}$, $M^*_{uv}(t) \preceq X^*(t)$ *where* $\preceq$ *denotes stochastic domination.*

From the previous lemma, we deduce the following.

LEMMA 5.21. *We have* $\mathbb{P}[M^*_{uv}(t) > \varepsilon_4 \log n, t < \tau_0' \wedge n^4] \leq \frac{1}{n^{10}}$.

PROOF. By Lemma 5.20, it suffices to prove the inequality in the statement with $M^*_{uv}(t)$ replaced by $X^*(t)$. Let $\mathcal{C}$ denote the following event:

$$\mathcal{C} = \{\forall \mathcal{T} \in [1, t], t' > \varepsilon_{16} n^2 \log n\ \#\{s \in [\mathcal{T}, \mathcal{T} + t'] : u(s) \neq v(s)\} \geq t'\varepsilon/4\}.$$

Then we have that for all $t - s > t' > \varepsilon_{16} n^2 \log n$,

$$\mathbb{E}(e^{\lambda X^s(t')} 1_{\mathcal{C} \cap \{t < \tau_0'\}}) \\ \le e^{\lambda K} \left(1 + \frac{9C_2}{n^2}(e^\lambda - 1)\right)^{t'(1-\varepsilon/4)} \\ \times \left(1 + \frac{9C_2}{n^2}(e^\lambda - 1) + \frac{K}{n^2}(1 - e^{-\lambda})\right)^{t'\varepsilon/4}.$$

Fix λ large enough such that $\lambda\varepsilon_4 > 20$. Choosing K sufficiently large depending on λ and ε, it follows that

$$\left(1 + \frac{9C_2}{n^2}(e^\lambda - 1)\right)^{1-\varepsilon/4} \left(1 + \frac{9C_2}{n^2}(e^\lambda - 1) + \frac{K}{n^2}(1 - e^{-\lambda})\right)^{\varepsilon/4} < 1$$

and hence

$$\mathbb{E}(e^{\lambda X^s(t')} 1_{\mathcal{C} \cap \{t < \tau_0'\}}) \le e^{\lambda K}.$$

By Markov's inequality, it follows that

$$\mathbb{P}(\{X^s(t') > \varepsilon_4 \log n\} \cap \mathcal{C} \cap \{t < \tau_0'\}) \le e^{-\lambda(K - \varepsilon_4 \log n)} \le \frac{1}{n^{19}}.$$

For $t' < \varepsilon_{16} n^2 \log n$, $X^s(t') - K$ is stochastically dominated by a $\mathrm{Bin}(\varepsilon_{16} n^2 \log n, \frac{9C_2}{n^2})$ variable. Using a Chernoff bound in this case, we get for n sufficiently large

$$\mathbb{P}[X^s(t') \ge \varepsilon_4 \log n] \le e^{-\varepsilon_4 \log n \log(\varepsilon_4 / 9C_2 \varepsilon_{16})/2} \le \frac{1}{n^{19}}$$

by choosing ε_{16} sufficiently small such that $\varepsilon_4 \log(\varepsilon_4 / 9C_2 \varepsilon_{16}) > 38$.

By taking a union bound over all s, t', it follows that

$$\mathbb{P}[X^*(t) > \varepsilon_4 \log n, \mathcal{C}, \{t < \tau_0'\}] \le \frac{1}{n^{11}}.$$

The result now follows from Lemma 5.19. □

We are now ready to prove Theorem 5.17.

PROOF OF THEOREM 5.17. The theorem follows from using Lemma 5.21, taking a union bound over all $(u, v) \in V^{(2)}$. □

5.4. *Degree estimate.* In this section, we consider the stronger stopping time τ_5 associated with the vertex degrees in $G(t)$ and prove the following theorem.

THEOREM 5.22. *We have for all $t \ge \delta n^2$, $\mathbb{P}[t + \delta n^2 \ge \tau_5, t + \delta n^2 < \tau_*, t < \tau] \le \frac{1}{n^{14}}$.*

The strategy to prove Theorem 5.22 is same as before. We shall show that whenever the degree of a vertex becomes too small or too large it gets a drift to the other side. To understand how we shall establish this, consider the example of a vertex degree becoming too large. We shall show that at most of the times a positive fraction of the edges incident to a given vertex v will be disagreeing, and hence if the degree becomes too large then the vertex will lose edges at a higher rate, thus providing a negative drift. We shall start with the following lemma.

LEMMA 5.23. *Let $\kappa_2 > 0$ be fixed. Let us condition on $\{\mathcal{F}_{t_1}, t_1 < \tau_0', N_*(t) = pn\}$. Let v be a fixed vertex in V. Let $X_v(t')$ denote the number of disagreeing edges incident to v at time t'. Then for sufficiently large C and sufficiently large $\beta = \beta(C)$, at time $t_2 = t_1 + \frac{Cn^2}{\beta}$ we have $\mathbb{P}[X_v(t_2) \notin (p(1-\kappa_2), 1-p(1-\kappa_2))D_v(t_2) \mid \mathcal{F}_{t_1}, \tau_0' > t_1] \le e^{-\varepsilon_{20} n}$ for some constant $\varepsilon_{20} > 0$.*

PROOF. The proof of this lemma goes along the same lines as that of Proposition 5.11. We shall therefore only give the sketch of the steps. Let the edges incident to v at time t_1 be $\{e_1, e_2, \dots, e_{D_v(t_1)}\}$. Let Y_i denote the indicator that the endpoints of e_i are disagreeing in $G(t_2)$. Notice that the number of edges adjacent to v that gets rewired in $\frac{Cn^2}{\beta}$ steps is $O(\frac{n}{\beta})$ with exponentially high probability, and hence arguing as in the proof of Proposition 5.11 it follows that the fraction of disagreeing edges incident to v at time t_2 is well approximated by $\frac{1}{D_v(t_1)} \sum_{i=1}^{D_v(t_1)} Y_i$ if β is sufficiently large. Hence, it suffices to prove that

$$\frac{1}{D_v(t_1)} \sum_{i=1}^{D_v(t_1)} Y_i \in \big(p(1-\kappa_2/2), 1-p(1-\kappa_2/2)\big)$$

with exponentially high probability. To this end, we choose a random subset of these edges of size $\varepsilon_{17} n$. Condition on a subset J of these edges of size at least $\varepsilon_{18} n = (1 - \frac{\kappa_2}{100})\varepsilon_{12} n$ which correspond to distinct bonds in $V^{(2)}$. Arguing as in the proof of Proposition 5.11, we see that it suffices to show that, conditionally,

$$\frac{1}{\varepsilon_{13} n} \sum_{j \in J} Y_j \in \big(p(1-\kappa_2/50), 1-p(1-\kappa_2/50)\big)$$

with exponentially high probability.

Consider the coupling described in Section 5.1 of the *rewire-to-random-** dynamics started with $G(t_1)$ with independent random walks started from v_j where e_j is placed in the bond (v, v_j) is $G(t_1)$. Let for $j \in J$ and $\tilde{\sigma} > \frac{C}{\beta}$, $U^0_{j,\tilde{\sigma}}$ and $U^1_{j,\tilde{\sigma}}$ denote the indicators that the position of the random walk started from v_j has the opinion 0 and opinion 1, respectively, at time $\tilde{\sigma}$. It follows by taking C sufficiently large that $\mathbb{E}(U^0_{j,\tilde{\sigma}}), \mathbb{E}(U^1_{j,\tilde{\sigma}}) \in (p(1-\kappa_2/100), 1-p(1-\kappa_2/100))$. Now the proof is completed arguing as in the proof of Proposition 5.11. $\square$

LEMMA 5.24. *With the notation as in Lemma 5.23, let A_t denote the event that for some t' with $t' \in [t+\frac{Cn^2}{\beta}, t+\delta n^2]$, $X_v(t') \notin (p-\varepsilon/20, 1-(p-\varepsilon/20))D_v(t')$. Then we have*

$$\mathbb{P}(A_t, t+\delta n^2 < \tau_0' \mid \mathcal{F}_t, t < \tau_0', N_*(t) = pn) \leq e^{-\varepsilon_{20}n/2}.$$

PROOF. This follows from a union bound and the previous lemma. □

LEMMA 5.25. *We have $\mathbb{P}[D_v(t+\delta n^2) \notin (\varepsilon/2, 1-\varepsilon/2)n, t+\delta n^2 < \tau_0' \mid \mathcal{F}_t, t < \tau_0, N_*(t) = pn] \leq e^{-\varepsilon_{21}n}$.*

PROOF. Let C_t denote the event that for all $t' \in [t+1, t+\delta n^2]$, $|N_*(t') - N_*(t)| \leq \frac{\varepsilon^2 n}{100}$. Let $H_t = A_t^c \cap C_t \cap \{t+\delta n^2 < \tau_0'\}$. Notice that on H_t, we have at time $t' \in [t, t+\delta n^2]$, the number of disagreeing edge at time t', denoted $Z(t')$ is in $(\frac{p(1-p)}{2} \pm \frac{\varepsilon^2}{20})n^2$. Also notice that on A_t, the number of disagreeing edges incident to v is at time t' is in $[p-\varepsilon/20, 1-(p-\varepsilon/20)]D_v(t)$. Set $X(t') = D_v(t') - (1-3\varepsilon/4)n$, also set $\Delta^* = (1-\frac{\beta}{n})$. It follows therefore that for $\lambda > 0$

$$\begin{aligned}
&\mathbb{E}(e^{\lambda X_{t'+1}} 1_{H_t} \mid \mathcal{F}_{t'}) \\
&\quad \leq e^{\lambda X_{t'}} \biggl(1 + (e^{\lambda} - 1)\frac{\Delta^* Z(t')}{N(n-1)} \\
&\qquad + (e^{-\lambda} - 1)\frac{\Delta^*(p-\varepsilon/8)(X_{t'} + (1-3\varepsilon/4)n)}{2N}\biggr).
\end{aligned}$$

Now take λ so small such that $e^{\lambda} - 1 \leq (1+\varepsilon^2/100)\lambda$ and $e^{-\lambda} - 1 \leq -(1-\varepsilon^2/100)\lambda$. Then we have

$$\begin{aligned}
&\mathbb{E}(e^{\lambda X_{t'+1}} 1_{H_t} \mid \mathcal{F}_{t'}) \\
&\quad \leq e^{\lambda X_{t'}} \biggl(1 + \lambda\Delta^*\biggl(\frac{(1+\varepsilon^2/50)Z_{t'}}{nN} - \frac{(p-\varepsilon/10)X_{t'}}{2N} \\
&\qquad - \frac{(p-\varepsilon/10)(1-3\varepsilon/4)n}{2N}\biggr)\biggr) \\
&\quad \leq e^{\lambda X_{t'}} \biggl(1 - \lambda\frac{\varepsilon X_{t'}}{8N}\biggr) \\
&\quad \leq e^{\lambda_* X_{t'}},
\end{aligned} \tag{5.12}$$

where $0 < \lambda_* < \lambda(1-\frac{\varepsilon}{10N})$ and since $p > \varepsilon$ implies that on H_t we have $\frac{(1+\varepsilon^2/50)Z_{t'}}{nN} < \frac{(p-\varepsilon/10)(1-3\varepsilon/4)n}{2N}$. It follows that there exists λ_0 bounded away from λ such that

$$\mathbb{E}(e^{\lambda X_{t+\delta n^2}} 1_{H_t} \mid \mathcal{F}_{t+Cn^2/\beta}) \leq e^{\lambda_0 X_{t+Cn^2/\beta}} 1_{H_t} \leq e^{\lambda_0 \varepsilon n/4},$$

that is,

$$\mathbb{E}\big(e^{\lambda X_{t+\delta n^2}} 1_{H_t} \mid \mathcal{F}_t, t < \tau_0\big) \le e^{\lambda_0 \varepsilon n/4}.$$

By Markov's inequality, it now follows that

$$\mathbb{P}\big[D_v\big(t+\delta n^2\big) \ge (1-\varepsilon/2)n, H_t \mid \mathcal{F}_t, t < \tau_0\big] \le e^{-(\lambda-\lambda_0)\varepsilon n/4}.$$

It follows from Lemma 5.24 and another random walk estimate that $\mathbb{P}(H_t^c, t + \delta n^2 < \tau_0')$ is exponentially small in n, which completes the proof of one side of the bound in this lemma.

The other side of the bound can be proved similarly by starting with λ negative and $X(t) = D_v(t) - 3\varepsilon/4$. This completes the proof of the lemma. $\square$

PROOF OF THEOREM 5.22. This theorem follows from Lemma 5.25 by taking a union bound over all vertices v, and all times $t' \in [t - \delta n^2, t]$ as in the proof of Theorem 5.16, and using Theorem 4.9. $\square$

5.5. *Multiple-edge estimates.* Finally, we consider the stronger stopping time τ_4 corresponding to the number of multiple edges adjacent to a fixed vertex in $G(t)$. Our main theorem in this subsection is the following.

THEOREM 5.26. *We have for all* $t \ge \delta n^2$, $\mathbb{P}[t + \delta n^2 \ge \tau_4, t + \delta n^2 < \tau_*, t < \tau] \le \frac{1}{n^4}$.

Theorem 5.26 will follow from the following proposition.

PROPOSITION 5.27. *Condition on* $\{\mathcal{F}_t, t < \tau_0\}$. *Let* v *be a vertex in* V. *Let* $X_{v,k}(t)$ *denote the number of vertices* $u \in G(t)$ *such that* $M_{uv}(t) \ge k$. *Then for each* $1 \le k \le 2\varepsilon_4 \log n$, *we have*

$$\mathbb{P}\big[X_{v,k}\big(t+\delta n^2\big) > C_1 10^{-k} n, t + \delta n^2 < \tau_0' \mid \mathcal{F}_t, t < \tau\big] \le \frac{1}{n^{10}}.$$

We postpone the proof of Proposition 5.27 for the moment and first show how this lemma implies Theorem 5.26.

PROOF OF THEROEM 5.26. For each fixed v, taking a union bound over all k, and then taking a union bound over all v, and then taking a union bound over times in $[t - \delta n^2, t]$ yields the theorem from Lemma 5.27. $\square$

Now we proceed with the proof of Proposition 5.27. For a fixed time t, by an r-edge in $G(t)$, we shall refer to a bond containing r edges at time t. To control the number of bonds with more than k edges adjacent to v at time $t + \delta n^2$ we do the following. We show first that for $\ell < k$, the fraction of bonds that contained ℓ

edges at time t but gains at least $(k-\ell)$ edges in the interval $[t, t+\delta n^2]$ is small with high probability. Next, we show that the number of k-edges at time t that gains at least one edges in the same time period is also small with high probability. Next, we show that a not too small fraction of the k-edges lose at least one edge in time $[t, t+\delta n^2]$ with high probability. Proposition 5.27 follows from the following three estimates.

LEMMA 5.28. *Condition on $\{\mathcal{F}_t, t<\tau_0\}$. Fix v and k as in Proposition 5.27. For $\ell = 0, 1, \dots, k-1$, recall $N(\ell \to k)$ is the number of ℓ-edge at time t that gained at least $(k-\ell)$ edges in time $[t, t+\delta n^2]$. Then we have for each $\ell \le (k-5)$:*

$$\mathbb{P}\big[N(\ell \to k) \ge \delta 5^{-(k-\ell)} C_1 10^{-k} n \mid \mathcal{F}_t, t<\tau_0\big] \le \frac{1}{n^{20}}.$$

Further, for $k-5<\ell<k$, we have

$$\mathbb{P}\big[N(\ell \to k) \ge 100\delta 10^{(k-\ell)} C_1 10^{-k} n \mid \mathcal{F}_t, t<\tau_0\big] \le \frac{1}{n^{20}}.$$

PROOF. This lemma is essentially obtained by reworking the proof of Lemma 4.6, and hence we give only a sketch. Condition on $\{\mathcal{F}_t, t<\tau_0\}$; fix v and k as in the statement of the lemma. Also fix $\ell<k$. Let $u_1, u_2, \dots, u_D$ be the vertices in V such that $\{M_{vu_i}(t)=\ell\}$. Now without loss of generality we assume $D = C_1 10^{-\ell} n$. Following the sequence of arguments in the proofs of Lemma 4.7, Lemma 4.8 and Lemma 4.6 (using the same notation as there), it follows that

$$\begin{aligned}&\mathbb{P}\big[N(\ell \to k) \ge \delta 5^{-(k-\ell)} C_1 10^{-k} n \mid \mathcal{F}_t, t<\tau_0\big]\\ &\quad \le e^{-4\delta C_2 D} + 2\mathbb{P}\left[\sum_{i=1}^{D} Z_i' \ge \delta 50^{-(k-\ell)} D\right],\end{aligned}$$

where $Z_i' = 1_{\{Y_i' \ge k-\ell\}}$ and $\{Y_i'\}_{1\le i\le D}$ a sequence of i.i.d. $\mathrm{Bin}(24\delta C_2 D, \frac{2}{D})$ variables. Using a Chernoff bound as before, it follows that if $k-\ell \ge 5$ then $\mathbb{P}[Z_i'=1] \le \delta 100^{-(k-\ell)}$ since $480000 C_2\delta < 1$. The first statement in the lemma now follows using a further Chernoff bound on $\sum Z_i'$.

For the case $k-5<\ell<k$, a simple first moment estimate gives an upper bound of $48C_2\delta$ on $\mathbb{P}[Z_i'=1]$. The second statement follows using another Chernoff bound and observing $C_2 \le 2$. □

LEMMA 5.29. *Assume the set-up of Lemma 5.28. Let $\widehat{N}(k)$ denote the number of bonds adjacent to v having at least k edges at time t that gained at least one edge during time $[t, t+\delta n^2]$. We then have*

$$\mathbb{P}\big[\widehat{N}(k) \ge 100\delta C_1 10^{-k} n\big] \le \frac{1}{n^{20}}.$$

PROOF. This proof is the same as the proof of the case $\ell = k-1$ case in Lemma 5.28; we omit the details. □

Now observe the following. Let $Y_{v,k}(t)$ denote the number of vertices $u \in G(t)$ such that $M_{uv}(t) = k$. In the set-up of Proposition 5.27, if $Y_{v,k}(t) < \frac{4}{5}C_1 10^{-k}n$, since $t < \tau_0$ it follows that $X_{v,k}(t) \le Y_{v,k}(t) + X_{v,k+1}(t) \le \frac{9}{10}C_1 10^{-k}n$. Clearly, since

$$X_{v,k}(t + \delta n^2) \le X_{v,k}(t) + \sum_{\ell=0}^{k-1} N(\ell \to k),$$

in this case Proposition 5.27 would follow from Lemma 5.28 by taking a union bound over all $\ell < k$. So, without loss of generality for the rest of this argument, we shall assume $Y_{v,k}(t) \ge \frac{4}{5}C_1 10^{-k}n$. We have the following lemma.

LEMMA 5.30. *Assume the set-up of Lemma 5.28. Let $k > \frac{10^{10}}{\varepsilon}$ and let $Y_{v,k}(t) \ge \frac{4}{5}C_1 10^{-k}n$. Also let $\widetilde{N}(k)$ denote the number of k-edges at time t which lost at least one edge in time $[t, t+\delta n^2]$. Then*

$$\mathbb{P}[\widetilde{N}(k) < 10^7 \delta C_1 10^{-k} n] \le \frac{1}{n^{20}}.$$

We defer the proof of Lemma 5.30 and instead first prove Proposition 5.27 using Lemmas 5.28, 5.29 and 5.30.

PROOF OF PROPOSITION 5.27. Condition on $\{\mathcal{F}_t, t < \tau_0\}$. Fix a vertex v. Now let C_1 be chosen sufficiently large depending (only) on ε, such that for all $k \le \frac{10^{10}}{\varepsilon}$ we have $C_1 10^{-k} > C_2$, and hence the proposition follows from Theorem 5.22. Now fix $k > \frac{10^{10}}{\varepsilon}$.

As argued above using Lemma 5.28, it suffices only to consider the case $Y_{v,k}(t) \ge \frac{4}{5}C_1 10^{-k}n$. Clearly, we have

$$X_{v,k}(t + \delta n^2) \le X_{v,k}(t) + \sum_{\ell=0}^{k-1} N(\ell \to k) + \widehat{N}(k) - \widetilde{N}(k).$$

Let $\mathcal{A}$ denote the event that for all $\ell \le k-5$, we have $N(\ell \to k) \le \delta 10^{-(k-\ell)} C_1 10^{-k} n$. Let $\mathcal{B}$ denote the event that for $k-5 < \ell < k$ we have $N(\ell \to k) \le 100\delta 10^{(k-\ell)} C_1 10^{-k} n$. Let $\mathcal{D}$ denote the event that $\widehat{N}(k) \le 100\delta C_1 10^{-k} n$. Finally, let $\mathcal{H}$ denote the event that $\widetilde{N}(k) \ge 10^7 \delta C_1 10^{-k} n$. Clearly, on $\mathcal{A} \cap \mathcal{B} \cap \mathcal{D} \cap \mathcal{H}$ we have

$$\sum_{\ell=0}^{k-1} N(\ell \to k) + \widehat{N}(k) - \widetilde{N}(k) \le 0$$

and since $X_{v,k}(t) \le C_1 10^{-k} n$ on $\{t < \tau_0\}$ it follows that $X_{v,k}(t+\delta n^2) \le C_1 10^{-k} n$. The proposition now follows from Lemmas 5.28, 5.29 and 5.30. $\square$

Coming back to the proof of Lemma 5.30 we shall need to show that a not too small fraction of the k-edges at time t loses at least one edge in time $[t, t+\delta n^2]$. For this particular argument, it will be convenient to work with the continuous time *rewire-to-random-** dynamics. First, we show that most k-edges spend a considerable fraction of the time with their endpoints disagreeing (clearly without this they cannot lose any edges). We have the following proposition.

PROPOSITION 5.31. *Let us condition on $\{\mathcal{F}_t, t < \tau_0\}$. Let v be a fixed vertex in V and let $2\varepsilon_4 \log n > k > \frac{10^{10}}{\varepsilon}$ be fixed. Let $u_1, u_2, \dots, u_D$ be the vertices in v such that we have $M_{vu_i}(t) = k$. Set $G(t) = H(0)$ and Run the following continuous time* rewire-to-random-* *process $H(\cdot)$ from time 0 to $\delta/2$. Each directed edge rings at rate 1. If the endpoints of the edge are agreeing in the current graph, no change occurs. If they are disagreeing, then we do a voter model step with probability $\frac{\beta}{n}$ and a rewire-to-random step with probability $(1-\frac{\beta}{n})$. Let Z_i be the indicator that (v, u_i) is disagreeing for less that $\frac{\varepsilon\delta}{200}$ time in $H(\cdot)$. Then we have $\mathbb{P}(\sum_{i=1}^{D} Z_i > \frac{\delta}{10} C_1 10^{-k} n, \tau_0' > \delta/2) \le e^{-\gamma\sqrt{n}}$ for some $\gamma > 0$.*

PROOF. Without loss of generality, we assume $D = \frac{4}{5} C_1 10^{-k} n \gg \sqrt{n}$. Let us choose a random subset $D^* \subseteq 1, 2, \dots, D$ with $|D^*| = \sqrt{n}$. It therefore suffices to prove that $\mathbb{P}(\sum_{i \in D^*} Z_i > \frac{\delta}{100}\sqrt{n} \mid D^*) \le e^{-\gamma\sqrt{n}}$. This fact is established by Lemma 5.32 and Lemma 5.33 below which completes the proof. $\square$

For the proof of the above proposition, we shall consider the coupling with the continuous time independent random walks on $G(t)$ described in Section 5.1. First, we show that for most of the vertices u_i the opinion of the random walkers started at u_i is different from the opinion of v in the continuous time *rewire-to-random-** dynamics for a significant amount of time. We have the following lemma.

LEMMA 5.32. *Condition on $\{\mathcal{F}_t, t < \tau_0\}$. Let us set $G(t) = H(0)$, and consider running the continuous time* rewire-to-random-* *dynamics $H(\cdot)$ from time 0 to $\delta/2$. Let $v, v_1, v_2, \dots, v_{\sqrt{n}}$ be fixed vertices in V. Let us consider the independent continuous time random walks described in Section 5.1. Let us $X_i^j(\cdot)$ be the random walk started from v_i on $H(2j\frac{C}{\beta})$ run for time $\frac{2C}{\beta}$. For $\frac{C}{\beta} \le s \le \frac{2C}{\beta}$, let $Y_i(2j\frac{C}{\beta} + s)$ is the indicator of the event that the opinion of $X_i^j(s)$ in $H(0)$ is different from the opinion of v in $H(s)$. Let $Y_i^* = \int_0^{\frac{\delta}{2}} Y_i(s)\,ds$. Further, let Z_i^* denote the indicator that $Y_i^* < \frac{\varepsilon\delta}{64}$. Then we have*

$$\mathbb{P}\Big[\sum_i Z_i^* \ge \frac{\delta}{1000}\sqrt{n}\Big] \le e^{-c\sqrt{n}}.$$

PROOF. For $h = 0, 2, \dots, \frac{C}{\theta} - 1$, $j = 1, 2, \dots, \frac{\delta\beta}{4C}$, let $\chi_v^{j,h} = 1$ if v spends majority of its time in the interval $[(2j-1)\frac{C}{\beta} + \frac{h\theta}{\beta}, (2j-1)\frac{C}{\beta} + \frac{(h+1)\theta}{\beta}]$ with opinion 1 and 0 otherwise. Let $\chi_i^{j,h} = 1$ if the opinion of $X_i^j(s) = 1$ for all $s \in [(2j-1)\frac{C}{\beta} + \frac{h\theta}{\beta}, (2j-1)\frac{C}{\beta} + \frac{(h+1)\theta}{\beta}]$, $\chi_i^{j,h} = 0$ if the opinion of $X_i^j(s) = 0$ for all $s \in [(2j-1)\frac{C}{\beta} + \frac{h\theta}{\beta}, (2j-1)\frac{C}{\beta} + \frac{(h+1)\theta}{\beta}]$, and $\chi_i^{j,h} = \star$ otherwise. Now let

$$U_i^{j,h} = 1_{\{\chi_i^{j,h}=1, \chi_v^{j,h}=0\}} + 1_{\{\chi_i^{j,h}=0, \chi_v^{j,h}=1\}}.$$

Let us fix h. Now choose θ sufficiently small so that the chance that the random walk takes a step in time θ/β is at most $\frac{\varepsilon}{4}$. Clearly, for a fixed realisation of the sequence $\chi_v^{j,h}$, and on $\{2(j-1)C/\beta < \tau_0'\}$, we have by Proposition 5.5, that $\mathbb{E}[U_i^{j,h} \mid \mathcal{F}_{2(j-1)C/\beta}] \ge \varepsilon/4$. Since the random walks are independent, it follows that

$$\mathbb{P}\left[\#\left\{i : \sum_j U_i^{j,h} \le \frac{\delta\beta\varepsilon}{8C}\right\} \ge e^{-\gamma'\beta}\sqrt{n}, \chi_v^{j,h}, \frac{\delta}{2} < \tau_0'\right] \le e^{-c\sqrt{n}}.$$

Taking a union bound over $3^{\delta\beta/4C}$ possible realisations of the sequence $\chi_v^{j,h}$ (for a fixed h), we get that

$$\mathbb{P}\left[\#\left\{i : \sum_j U_i^{j,h} \le \frac{\delta\beta\varepsilon}{32C}\right\} \ge e^{-\gamma'\beta}\sqrt{n}, \frac{\delta}{2} < \tau_0'\right] \le e^{-c\sqrt{n}}$$

for some constant γ' and $c > 0$.

Now taking a union bound over h, we get

$$\mathbb{P}\left[\#\left\{i : \sum_j \sum_h U_i^{j,h} \le \frac{\delta\beta\varepsilon}{32\theta}\right\} \ge \frac{C}{\theta} e^{-\gamma'\beta}\sqrt{n}, \frac{\delta}{2} < \tau_0'\right] \le \frac{C}{\theta} e^{-c\sqrt{n}}.$$

Now notice that on $\{\sum_j \sum_h U_i^{j,h} > \frac{\delta\beta\varepsilon}{32\theta}\}$, we have $Y_i^* > \frac{\varepsilon\delta}{64}$, and the proof of the lemma is completed by taking β sufficiently large. □

To complete the proof of Proposition 5.31, we still need to show that approximating the opinions of the vertices u_i by the opinion of the random walkers via the above coupling is good enough for our purposes. This is the content of the next lemma.

LEMMA 5.33. *Assume the setting of Lemma 5.32. Also let $\tilde{Y}_i$ denote the amount of time the bond (v, v_i) is disagreeing in $[0, \frac{\delta}{2}]$. Then there is a coupling of the continuous time evolving voter model with the continuous time random walks started at v_i as described in Lemma 5.32 such that*

$$\mathbb{P}\left[\#\left\{i : \tilde{Y}_i \le Y_i^* - \frac{\varepsilon\delta}{128}\right\} \ge \frac{\delta}{1000}\sqrt{n}, \frac{\delta}{2} < \tau_0'\right] \le e^{-c\sqrt{n}}$$

for some constant c.

PROOF. Consider the coupling described in Section 5.1, with the obvious modification for the continuous time dynamics. Define $Y_{i,j}=0$ if the opinion of $X_i^j(s)$ is the same as the opinion of v_i in $H(\frac{2jC}{\beta}+s)$ for all $s\in[0,\frac{2C}{\beta}]$ and $Y_{i,j}=1$ otherwise. There exists a function $g(\beta)\ll\beta$ such that $g(\beta)\to\infty$ as $\beta\to\infty$ for which we have

$$\mathbb{P}\left[\sum_{j=1}^{\delta\beta/4C}\sum_{i=1}^{\sqrt{n}}Y_{i,j}\geq 10C\delta g(\beta)\sqrt{n},\frac{\delta}{2}<\tau_0'\right]\leq e^{-c\sqrt{n}} \tag{5.13}$$

for some constant $c>0$. Indeed we can prove (5.13) by following the sequence of arguments in Lemma 5.6, Lemma 5.7, Lemma 5.8 and Lemma 5.9. We omit the proof. Again notice that, $\tilde{Y}_i-Y_i^*\geq-(\sum_j Y_{i,j})\frac{2C}{\beta}$. It follows that

$$\mathbb{P}\left[\#\left\{i:\tilde{Y}_i\leq Y_i^*-\frac{\varepsilon\delta}{128}\right\}\geq\frac{2560C^2g(\beta)}{\beta\varepsilon\delta}\sqrt{n},\frac{\delta}{2}<\tau_0'\right]\leq e^{-c\sqrt{n}}.$$

The proof of the lemma is completed by taking β sufficiently large. $\square$

Now that we have established Proposition 5.31, we still need to show that if most of the bonds (v,u_i) are disagreeing for a significant time then a not too small fraction of them loses at least one edge in time $[0,\frac{\delta}{2}]$ with high probability.

LEMMA 5.34. *Condition on $\{\mathcal{F}_t,t<\tau_0\}$. Assume the setting of Proposition 5.31 and let v,k be fixed as there. Also assume $D\geq\frac{4}{5}C_1 10^{-k}n$. Let W_i be the indicator that the bond (v,u_i) loses at least one edge by time $\frac{\delta}{2}$. Then we have*

$$\mathbb{P}\left[\sum_{i=1}^{D}W_i<10^7\delta C_1 10^{-k}n,\tau_0'>\frac{\delta}{2}\right]\leq\frac{1}{n^{20}}.$$

PROOF. Without loss of generality, assume $D=\frac{4}{5}C_1 10^{-r}n>>\sqrt{n}$. In the continuous time model, the rate at which a bond with k disagreeing edges lose an edge is $k(1-\frac{1}{n})$, and the rings of the different edges are independent. Let $S_1,S_2,\dots,S_D$ be the number of edges lost by the bonds $vu_1,vu_2,\dots,vu_D$, respectively. Let $\{S_i'\}_{1\leq i\leq D}$ be a collection of independent Poi$(4\times10^7\delta)$ variables. Since $\frac{k\delta\epsilon}{200}>5\times10^7\delta$, it follows from Proposition 5.31 that there exist a coupling such that

$$\mathbb{P}\left[\#\{i:S_i\leq S_i'\}\geq\frac{\delta}{10}C_1 10^{-k}n,\tau_0'>\frac{\delta}{2}\right]\leq e^{-c\sqrt{n}}$$

for some constant $c>0$. Also observe that $\mathbb{P}[S_i'\geq1]\geq2\times10^7\delta$ (recall $\delta<10^{-10}$), and hence by a large deviation estimate it follows that

$$\mathbb{P}[\#\{i:S_i'\geq1\}<1.5\times10^7\delta C_1 10^{-k}n]\leq\frac{1}{n^{25}}.$$

The lemma follows combining the above estimates. □

Finally, we are ready to prove Lemma 5.30.

PROOF OF LEMMA 5.30. Without loss of generality, we assume $Y_{v,k}(t) = \frac{4}{5}C_1 10^{-k}n$. The proof now follows from Lemma 5.34 and the obvious coupling of the continuous time rewire-to-random-* dynamics with the discrete time rewire-to-random-* dynamics and observing that with exponentially high probability, the number of steps taken in the discrete time process up to time $\frac{\delta}{2}$ in the continuous time process is less that δn^2. □

5.6. *Completing the proof of Theorem* 3.2. Now we are ready to prove Theorem 3.2.

PROOF OF THEOREM 3.2. Using a random walk estimate, it is clear that $\mathbb{P}[\tau_0 > n^4] = o(1)$. Also it is clear from the properties of an Erdős–Rényi graph that $\mathbb{P}[\tau_0 < \delta n^2] = o(1)$. Now for $k \geq 0$, and $i = 2, 3, 4, 5$, let $A_{k,i}$ denote the event $\{k\delta n^2 < \tau_0, (k+1)\delta n^2 < \tau_*, (k+1)\delta n^2 \geq \tau_i\}$. Using Theorem 5.16, Theorem 5.17, Theorem 5.26 and Theorem 5.22 and taking a union over $0 \leq k \leq n^2/\delta$, it follows that

$$\mathbb{P}\big[\tau_0 < \tau_* - \delta n^2\big] \leq o(1) + \sum_{i,k} \mathbb{P}(A_{k,i}) = o(1).$$

This completes the proof. □

6. Rewire-to-random eventually splits. In this section, we prove Theorem 2. For this section, we shall consider running the *rewire-to-random* model with a different initial condition. For $0 < p < 1$, let $\mathcal{G}^*(p)$ be the following subset of the state space of our Markov chain, that is, let $\mathcal{G}^*(p)$ is a set of multi-graphs of n vertices with labelled edges where each vertex has either of the two opinions 0 and 1, such that $N_1(G) = pn$ and the number of edges in G is in $[\frac{12n^2}{50}, \frac{13n^2}{50}]$.

Recall the stopping times $\tau = \min\{t : \mathcal{E}^\times(t) = \varnothing\}$ for the process to terminate and for $\eta \in (0, \frac{1}{2}]$, recall $\tau_*(\eta) = \min\{t : N_*(t) \leq \eta n\}$ is the first time the minority opinion density reaches η. Denote $\tilde{\tau} = \tilde{\tau}(p) = \tau_*(p/2)$. Theorem 2 will follow from the next theorem.

THEOREM 6.1. *Let $\beta > 0$ be fixed. Consider running the* rewire-to-random *model with relabelling rate β starting with the state $G(0)$. Then there exists $p = p(\beta)$ sufficiently small such that for all $G(0) \in \mathcal{G}^*(p)$, we have $\tau < \tilde{\tau}$ with high probability.*

We postpone the proof of Theorem 6.1 and instead show first how this implies Theorem 2.

PROOF OF THEOREM 2. Since in the set up of this theorem $G(0)$ is distributed as $G(n, \frac{1}{2})$, it follows that with high probability the number of edges in $G(0)$ is in $[\frac{12n^2}{50}, \frac{13n^2}{50}]$. Let $p = p(\beta)$ be sufficiently small so that the conclusion of Theorem 6.1 holds. It follows that on $\{\tau_*(p) < \infty\}$, with high probability $G(\tau_*(p)) \in \mathcal{G}^*(p)$. Since the *rewire-to-random* dynamics is symmetric in the opinions 0 and 1, it follows from Theorem 6.1 that $\tau < \tilde{\tau}$ with high probability. This completes the proof of the theorem. □

Before starting with the proof of Theorem 6.1 we make the following definitions. Let us fix $G(0) \in \mathcal{G}^*(p)$. Let S be the set of vertices in $G(0)$ with degree at most $10n$ and let T be the set of vertices with degree more than $10n$. Clearly, $|S| \geq \frac{24n}{25}$. We next describe an equivalent way of constructing the *rewire-to-random* dynamics started with $G(0)$.

An equivalent construction of the dynamics:

Let $\{X_i\}$ and $\{X_i'\}$ be two sequences of i.i.d. Geom($\frac{\beta}{n}$) variables (taking values in $\{0, 1, \dots\}$). Let $\{Z_i\}$ be a sequence of i.i.d. Ber($\frac{1}{2}$) variables. Let $\{W_i\}$ be a sequence of vertices of G where each W_i is a uniformly chosen vertex of G. All these sequences are distributed independently of each other.

We now describe how to run the process starting with $G(0)$ using only the randomness in the above sequences and the randomness used to choose a disagreeing edge uniformly at each step. Having chosen a disagreeing edge the variables Z_i will be used to designate one of the endpoints of the edge uniformly as the root of the current (relabelling or rewiring) update. Also for each vertex v in V, we shall define a sequence $K_i(v)$. To start with, list the vertices in V in some order, say $\{v_1, v_2, \dots, v_n\}$. Define $K_0(v_j) = X_j'$ for all j. This encodes the number of updates at the vertex v_j (i.e., the number of moves with v_j being the root) before it changes its opinion for the first time which clearly has a Geom($\frac{\beta}{n}$) distribution. Roughly speaking, for each vertex v the sequence $\{K_i(v)\}$ will be a counter which shall denote how many more rewiring updates one needs to make at v before the next relabelling update. Once the counter runs to 0, the next update at that vertex is a relabelling one, and a new value from either the sequence $\{X_i\}$ or the sequence $\{X_i'\}$ will be assigned to the counter. We describe the process formally below.

We shall define the sequences L_i, L_i', T_i recursively; these will be indices of different elements chosen from $\{X_i\}$, $\{X_i'\}$ and $\{W_i\}$, respectively. Let $L_0 = T_0 = 0$ and $L_0' = n$. At step i, pick a disagreeing edge e uniformly at random, if such an edge exists. If $Z_i = 1$, then choose the vertex with opinion 1 to be the root of the rewiring or relabelling step, if $Z_i = 0$, choose the other one. Let v be the chosen vertex. If v is in S and the opinion of v is 0, do the following. Set $L_i' = L_{i-1}'$. If $K_{i-1}(v)$ is positive, then define $T_i = \min\{k > T_{i-1} : W_k \neq v\}$, that is T_i is the index of the first hitherto uninspected element in $\{W_j\}$ which allows a legal rewiring move. Rewire the edge to W_k and reduce $K_i(v)$ by 1, and set $L_i = L_{i-1}$. If $K_{i-1}(v) = 0$, then relabel v and set $L_i = L_{i-1} + 1$ and $K_i(v) = X_{L_i}$, in this case, also set $T_i = T_{i-1}$. If v is not in S, or the opinion of v is 1, then do the same as in the previous case except use elements from the sequence X' and L' in stead of the elements from sequence X and L, and change the values in the sequence L' instead of the sequence L. It is easy to see that this is indeed an implementation of the *rewire-to-random* dynamics.

With this implementation let us consider running the process for $10n^2$ steps if possible, that is, if it does not reach τ. We shall need the following sequence of lemmas. The first follows immediately from the fact that the number of vertices with opinion 1 does a random walk until time τ.

LEMMA 6.2. *For a fixed $p > 0$ and $G(0) \in \mathcal{G}^*(p)$, the number of vertices of opinion 1 remains between $pn/2$ and $3pn/2$ throughout the first $10n^2$ steps w.h.p.*

We call an element of the sequence X *stubborn* if it is at least $25n$.

LEMMA 6.3. *Let $Y = \#\{i \le L_{10n^2} : X_i > 25n\}$ denote the number of stubborn elements of X which are used in first $10n^2$ steps. Then with high probability, $N_1(10n^2) \ge Y$, that is, the number of vertices with label 1 after $10n^2$ steps is at least the number of used* stubborn *elements of the sequence X.*

PROOF. Let $\mathcal{S}$ denote the following event:

$$\mathcal{S} = \{\forall v \in G : \#\{i \le T_{10n^2} : W_i = v\} \le 14n\}.$$

We show that on $\mathcal{S}$, the vertices in S that each stubborn element gets assigned to [i.e., those v such that $X_\ell = K_i(v)$ for some i and some stubborn X_ℓ] are distinct and each of them has label 1 after $10n^2$ steps. Consider a specific stubborn element, suppose it was used and assigned to the vertex v. By definition, at that point the opinion of v was 1. Now by definition of *stubbornness*, before it changes its opinion again, v needs to be the root of at least $25n$ rewiring moves. Notice now that the number of rewirings rooted at v is at most the sum of the initial degree of v (which is at most $10n$ since $v \in S$) and the number of rewirings to v (which is at most $14n$ on $\mathcal{S}$). Hence, the vertex v never changes its opinion again and in particular is never associated with any other stubborn element. Hence, corresponding to each used stubborn element, there are distinct vertices in V which have opinion 1 after $10n^2$ steps.

It remains to show that $\mathcal{S}$ occurs with high probability. First, notice that using an argument similar to that used in the proof of Lemma 2.2, it follows that $\mathbb{P}(T_{10n^2} > 11n^2)$ is exponentially small in n. Also, we note that for each $v \in V$, the chance that v occurs more than $14n$ times in the first $11n^2$ elements of the list W is exponentially small in n using a Chernoff bound. Taking a union bound over all the vertices completes the proof of the lemma. $\square$

LEMMA 6.4. *Let RL_{SS} denote the number of times (up to $10n^2$ steps) a relabelling occurs when an edge with both endpoints in S was chosen. Then, for p sufficiently small, $RL_{SS} \leq \frac{\beta n}{20}$ w.h.p.*

PROOF. Let RL_{SS}^+ denote the number of times (up to $10n^2$ steps) we have a relabelling changing an opinion to 1 after an edge with both endpoints in S was chosen. Notice that each time we choose an edge with both endpoints in S, and do a relabelling update changing an opinion from 0 to 1, and element from the sequence X gets used. Hence, it follows from Lemma 6.2 and Lemma 6.3 it follows that w.h.p.

$$RL_{SS}^+ \leq \min_i \left\{ \#\{j \leq i : X_j \geq 25n\} > \frac{3pn}{2} \right\}.$$

Since each X_j is a $\mathrm{Geom}(\frac{\beta}{n})$ variable it follows that $\mathbb{P}(X_j \geq 25n) = (1 - \frac{\beta}{n})^{25n} \geq e^{-50\beta}$. Since X_j's are i.i.d., it follows that for p sufficiently small (depending on β), within first $\frac{\beta n}{50}$ elements of X, there are more than $\frac{3pn}{2}$ *stubborn* elements with high probability. It follows that with high probability $RL_{SS}^+ \leq \frac{\beta n}{50}$.

Now notice that each time a relabelling occurs when an edge with both endpoints in S is chosen, with probability $\frac{1}{2}$ the relabelling changes an opinion to 1 and these events are independent of one another. It follows that $\mathbb{P}(RL_{SS} > \frac{\beta n}{20}, RL_{SS}^+ \leq \frac{\beta n}{50})$ is exponentially small in n. It now follows that for p sufficiently small, $RL_{SS} \leq \frac{\beta n}{20}$ w.h.p. □

LEMMA 6.5. *Let R_{SS} be the number of times up to $10n^2$ an edge with both endpoints on S was picked. For p sufficiently small, $R_{SS} \leq \frac{n^2}{10}$ w.h.p.*

PROOF. Each time an edge is picked, it leads to a relabelling with probability $\frac{\beta}{n}$. It follows using a Chernoff bound that

$$\begin{aligned}\mathbb{P}\left(R_{SS} > \frac{n^2}{10}, RL_{SS} \leq \frac{\beta n}{20}\right) &\leq \mathbb{P}\left(\mathrm{Bin}\left(\frac{n^2}{10}, \frac{\beta}{n}\right) \leq \frac{\beta n}{20}\right) \\ &\leq e^{-\frac{\beta n}{80}}.\end{aligned}$$

The lemma now follows from Lemma 6.4. □

Let W_{ST} denote the total number of rewiring moves (up to time $10n^2$) where a disagreeing edge with one endpoint in S and another endpoint in T is rewired. Let W_{SS} and W_{TT} be defined similarly. Let Y_{SS} denote the number of edges with both endpoints in S at the end of the process (i.e., after running $10n^2$ steps). Y_{ST}, Y_{TT} are defined similarly. Finally, let R_{ST} (resp., R_{TT}) be the number of times (up to $10n^2$) a disagreeing edge is picked with one endpoint in S and another in T (resp., both endpoints in T). We have the following lemmas.

LEMMA 6.6. *For p sufficiently small, $R_{ST} \le 3n^2$ w.h.p.*

PROOF. Each time an edge with one endpoint in S and the other in T is rewired, with probability $\frac{1}{2}$ it is rewired with the root at the vertex in S and independent of that with probability at least $\frac{|S|-1}{n-1}$ the edge is rewired to a vertex in S. That is, after rewiring an edge with one endpoint in S and another in T, the chance that it becomes an edge with both endpoints in S is at least $\frac{|S|-1}{2(n-1)} \ge \frac{23}{50}$ for n sufficiently large. Let $W_{ST\to SS}$ denote the number of such rewirings. It follows that $\mathbb{P}(W_{ST\to SS} \le \frac{11n^2}{10}, W_{ST} > \frac{5n^2}{2})$ is exponentially small in n^2. Now notice that

$$n^2 \ge Y_{SS} \ge -R_{SS} + W_{ST\to SS}$$

and it follows from Lemma 6.5 that for p sufficiently small $W_{ST\to SS} \le \frac{11n^2}{10}$ w.h.p., and hence $W_{ST} \le \frac{5n^2}{2}$ w.h.p. Since each time a disagreeing edge is picked, with probability $1 - \frac{\beta}{n}$, it leads to a rewiring, it follows that $\mathbb{P}(R_{ST} > 3n^2, W_{ST} \le \frac{5n^2}{2})$ is exponentially small in n^2, and hence for p sufficiently small $R_{ST} \le 3n^2$ w.h.p. □

LEMMA 6.7. *For p sufficiently small $R_{TT} \le 6n^2$ w.h.p.*

PROOF. Arguing as in the proof of Lemma 6.6, we have that after rewiring an edge with both endpoints in T, the chance that it becomes an edge with one endpoint in S and another in T is $\frac{|S|}{n} \ge \frac{24}{25}$. Let $W_{TT\to ST}$ denote the number of such rewirings. It follows that $\mathbb{P}(W_{TT\to ST} \le 4n^2, W_{TT} > 5n^2)$ is exponentially small in n^2. Now notice that

$$n^2 \ge Y_{ST} \ge -R_{ST} + W_{TT\to ST}$$

and it follows from Lemma 6.6 that for p sufficiently small $W_{TT\to ST} \le 4n^2$ w.h.p. It follows that $W_{TT} \le 5n^2$ w.h.p. Arguing as in the proof of Lemma 6.6 we conclude that $R_{TT} \le 6n^2$ w.h.p. □

We are now ready to prove Theorem 6.1.

PROOF OF THEOREM 6.1. From Lemma 6.5, Lemma 6.6 and Lemma 6.7 we have that for $p = p(\beta)$ sufficiently small and $G(0) \in \mathcal{G}^*(p)$, we have $R_{SS} + R_{ST} + R_{TT} < 10n^2$ with high probability. This implies after $10n^2$ steps there are no disagreeing edge in the graph, that is, for p sufficiently small $\tau \le 10n^2$ w.h.p. The theorem now follows from Lemma 6.2. □

7. Modifications for the rewire-to-same model. We can prove Theorem 1 for the rewire-to-same model in a similar manner. To avoid repetitions, we only point out the main differences here. Notice that, with respect to our proof in previous sections, the major difference between the two dynamics that causes some inconvenience is that in the *rewire-to-same* dynamics a vertex with minority opinion is likely to receive edges at a higher rate than a vertex with majority opinion. But as long as the minority opinion density does not become too small, the difference is of a bounded factor, and it turns out that the arguments can be modified to accommodate this. We now point out the main lemmas from the previous sections that need to be modified for the *rewire-to-same* dynamics.

7.1. *Small β case.* Notice that the only place that needs a modification is Lemma 2.2. One needs to define for this case L_{i+1} as the first entry after L_i to which a rewiring move is legal. Here, it is not true that $L_{6n^2} \leq 13n^2/2$ with exponentially high probability. But notice that, on $t < \tau_*(1/3)$, $L_{i+1} - L_i$ is stochastically dominated by a Geom($\frac{1}{3}$) variable, and hence, one can say $L_{6n^2} < 20n^2$ with exponentially large probability. The rest of the proof is in essence same up to some minor changes in constants.

7.2. *Large β case.* The proof in the large β case also follows along the similar lines. Instead of the *rewire-to-same* dynamics, we consider the *rewire-to-same-** dynamics, where instead of a disagreeing edge, at each step we pick an edge at random, and do not do anything if the edge happens to be agreeing. Most of necessary changes occur while bounding the number of incoming edges to a vertex in time $[t, t+\delta n^2]$. But on $\tau < \tau_*(\varepsilon)$, one can bound the number of incoming edges to a vertex v in that time by a Bin(δn^2, $\frac{1}{\varepsilon n}$) variable.

The only other place where a somewhat significant modification is necessary is in the bound for large cuts. Instead of Proposition 5.11 and Proposition 5.13, one needs to show that for any given cut S and T, with $|S| \wedge |T| \geq \varepsilon_2 n$, the number of edges with one endpoint in S and another endpoint in T, such that the S end point has opinion 0 and the T endpoint has opinion 1, is roughly about $p(1-p)$ fraction of the total number of edges with exponentially high probability, and similar bounds on other similar quantities. It can be checked that all these can be obtained following a similar line of arguments as in the proof of Proposition 5.11. We omit the details. The rest of the bounds can then be obtained by suitably modifying the martingale calculations in Proposition 5.15. The whole proof can then be carried out with some minor changes of constants.

8. Open questions. While we establish some rigorous results for the evolving voter model on dense random graphs, the picture is far from clear. We conclude with the following natural questions, that are still open:

- *What happens eventually in the* rewire-to-same *model?* Notice that we do not have any result analogous to Theorem 2 in the *rewire-to-same* model. It is a natural question to ask whether in the rewire-to-same model, is there a positive fraction of both opinions present when the process reaches an absorbing state? As we have mentioned before, in [4] it was conjectured that, for the sparse graphs (with constant average degree), in the rewire-to-same model, one of the opinions take over almost the whole graph, but it is not known rigorously.
- *Is there a sharp transition in β?* Another natural question to ask is if there is any value β_0 such that for $\beta < \beta_0$ we have behaviour as in Theorem 1(i) and for $\beta > \beta_0$, we have behaviour as in Theorem 1(ii)?
- *What can we say about sparser graphs?* We can prove by an argument similar to proof of Theorem 2. for sparser graph with suitably rescaled relabelling rate [i.e., $G(0) \sim G(n, \frac{\lambda}{n})$] that there exists $\beta_0 > 0$, such that for all $\beta < \beta_0$, with high probability the process stops before the density of the opinions change. But the other side of the phase transition seems harder to prove. The main difficulty seems to be the presence of a few vertices of very high degree. It is another interesting question to investigate whether one could prove results about evolving voter models starting with not too sparse graphs, for example, $G(0) \sim G(n, n^{\alpha-1})$ for some $0 < \alpha < 1$? Following our arguments, Basak, Durrett and Zhang [2] very recently observed that analogues of Theorem 1(i) and Theorem 2 holds in this case.

Acknowledgements. The authors would like to thank Rick Durrett for introducing them to the model and an anonymous referee for helpful suggestions that helped improve the exposition of the paper. This work was completed while Riddhipratim Basu was a graduate student at the Department of Statistics, University of California, Berkeley.

REFERENCES

[1] ALDOUS, D. and FILL, J. A. (2002). *Reversible Markov Chains And Random Walks On Graphs*. Unfinished monograph. Recompiled 2014. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] BASAK, A. DURRETT, R. and ZHANG, Y. (2015). The evolving voter model on thick graphs. Preprint. Available at arXiv:1512.07871.

[3] DURRETT, R. (2010). *Random Graph Dynamics*. Cambridge Univ. Press, Cambridge. MR2656427

[4] DURRETT, R., GLEESON, J. P., LLOYD, A. L., MUCHA, P. J., SHI, F., SIVAKOFF, D., SOCOLAR, J. E. S. and VARGHESE, C. (2012). Graph fission in an evolving voter model. *Proc. Natl. Acad. Sci. USA* **109** 3682–3687. MR2903372

[5] FRIEDLAND, S. and NABBEN, R. (2002). On Cheeger-type inequalities for weighted graphs. *J. Graph Theory* **41** 1–17. MR1919163

[6] GIL, S., ZANETTE, D. H., BARILOCHE, C. A. and NEGRO, R. (2006). Coevolution of agents and networks: Opinion spreading and community disconnection. *Phys. Lett. A* **356** 89–94.

[7] GROSS, T. and BLASIUS, B. (2008). Adaptive coevolutionary networks: A review. *J. R. Soc. Interface* **5** 259–271.

[8] HENRY, A. D., PRAŁAT, P. and ZHANG, C.-Q. (2011). Emergence of segregation in evolving social networks. *Proc. Natl. Acad. Sci. USA* **108** 8605–8610. MR2813308

[9] HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** 643–663. MR0402985

[10] HOLME, P. and NEWMAN, M. (2006). Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74** 056108.

[11] KIMURA, D. and HAYAKAWA, Y. (2008). Coevolutionary networks with homophily and heterophily. *Phys. Rev. E* **78** 016103.

[12] KOZMA, B. and BARRAT, A. (2008). Consensus formation on coevolving networks: Groups' formation and structure. *J. Phys. A* **41** 224020. MR2453832

[13] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR2466937

[14] LIGGETT, T. M. (1985). *Interacting Particle Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **276**. Springer, New York. MR0776231

[15] MALIK, N. and MUCHA, P. J. (2013). Role of social environment and social clustering in spread of opinions in coevolving networks. *Chaos* **23** 043123. MR3389752

[16] NEWMAN, M. E. J. (2010). *Networks*. *An Introduction*. Oxford Univ. Press, Oxford. MR2676073

[17] ROGERS, T. and GROSS, T. (2013). Consensus time and conformity in the adaptive voter model. *Phys. Rev. E* **88** 030102.

[18] SHI, F., MUCHA, P. J. and DURRETT, R. (2013). Multiopinion coevolving voter model with infinitely many phase transitions. *Phys. Rev. E* **88** 062818.

[19] SKYRMS, B. and PEMANTLE, R. (2000). A dynamic model of social network formation. *Proc. Natl. Acad. Sci. USA* **97** 9340–9346.

[20] SOOD, V. and REDNER, S. (2005). Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94**.

[21] SUCHECKI, K., EGUÍLUZ, V. M. and SAN MIGUEL, M. (2005). Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution. *Phys. Rev. E* **72** 036132.

[22] VAZQUEZ, F. (2013). Opinion dynamics on coevolving networks. In *Dynamics on and of Complex Networks*. *Vol*. 2. *Model*. *Simul*. *Sci*. *Eng*. *Technol*. (A. Mukherjee, M. Choudhury, F. Peruani, N. Ganguly and B. Mitra, eds.) 89–107. Birkhäuser/Springer, New York. MR3156593

[23] VAZQUEZ, F., EGUÍLUZ, V. M. and MIGUEL, M. S. (2008). Generic absorbing transition in coevolution dynamics. *Phys. Rev. Lett.* **100** 108702.

[24] VOLZ, E. and MEYERS, L. A. (2007). Susceptible-infected recovered epidemics in dynamic contact networks. *Proc. Biol. Sci.* **274** 2925–2933.

[25] VOLZ, E. and MEYERS, L. A. (2009). Epidemic thresholds in dynamic contact networks. *J. R. Soc. Interface* **6** 233–241.

[26] YI, S. D., BAEK, S. K., ZHU, C.-P. and KIM, B. J. (2013). Phase transition in a coevolving network of conformist and contrarian voters. *Phys. Rev. E* **87** 012806.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: rbasu@stanford.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CALIFORNIA 94720
USA
E-MAIL: sly@stat.berkeley.edu