Abstract. We determine the large size limit of a network of interacting Hawkes Processes on an adaptive network. The flipping of the node variables is taken to have an intensity given by the mean-field of the afferent edges and nodes. The flipping of the edge variables is a function of the afferent node variables. The edge variables can be either symmetric or asymmetric. This model is motivated by applications in sociology, neuroscience and epidemiology. In general, the limiting probability law can be expressed as a fixed point of a self-consistent Poisson Process with intensity function that is (i) delayed and (ii) depends on its own probability law. In the particular case that the edge flipping is only determined by the state of the pre-synaptic neuron (as in neuroscience) it is proved that one obtains an autonomous neural-field type equation for the dual evolution of the synaptic potentiation and neural potentiation.

THE HYDRODYNAMIC LIMIT OF HAWKES PROCESSES ON ADAPTIVE STOCHASTIC NETWORKS

JAMES MACLAURIN *

We study stochastic processes on Adaptive Networks, consisting of node variables and edge variables that evolve stochastically in time [12]. Both the node and edge variables take values in a finite state space, with Poissonian jumps between states. The intensity of the edge variable flipping is a function of the states of the nodes at either end, and the flipping of the node variables is a 'mean-field' of the states of all edges and accompanying nodes. There are many applications of this sort of network [11]: including neuroscience (the edge dynamics corresponds to slow synaptic dynamics and learning), epidemiology (for example, if individuals are more likely to self-isolate if they get infected).

Despite the many many applications, there does not seem to exist a general 'Mckean-Vlasov' type equation that yields the large n limiting dynamics. Previous work, including by this author, only obtained limiting equations by resorting to implicit delay-stochastic differential equations [15]. There are several studies of related systems, including [10], Many scholars have also considered the hydrodynamic limit of large networks of interacting neurons on inhomogeneous graphs, including [7, 14, 1, 5, 6, 3].

A very important application of these results is that it yields a general formalism for generating random graphs, complementing for instance [13]. If, for example, the empirical measure concentrates at a single value in the large n limit, then it will automatically yield an accurate understanding of the local structure of the graph. One must also note that there exists a variety of adaptive network models which are not of the mean-field type, such as the adaptive voter model formulated by Durrett [4].

0.1. Notation. The index set of the nodes is $I_n := \{1, 2, \ldots, n\}$. Let Γ and Γ_E be discrete sets, specifying the possible states of the node variables and edge variables. Let $\mathcal{D}([0,T],\Gamma)$ and $\mathcal{D}([0,T],\Gamma_E)$ specify the set of all cadlag trajectories taking values in (respectively) Γ and Γ_E. This means that any $x \in \mathcal{D}([0,T],\Gamma)$ must be (i) piecewise constant, (ii) with only a finite number of discontinuities, (iii) possessing left limits and continuous from the right.

Let $\mathcal{X}_T \subset \mathcal{D}([0,T],\mathbb{R})$ consist of all cadlag functions that are (i) non-decreasing, and (ii) equal to 0 at time 0. For any $x, y \in \mathcal{D}([0,T],\Gamma)$, define

$$d_t(x,y) = \int_0^t \chi\{x_s \neq y_s\} ds. \tag{0.1}$$

We note that d_t generates the Skorohod Toplogy on $\mathcal{D}([0,T],\Gamma)$, however d_t is not a complete metric. For a positive integer a, let $\mathcal{Z}_{T,a}$ consist of all $x \in \mathcal{D}([0,T],\Gamma)$ such that the number of points of discontinuity is less than or equal to a. One can check that $\mathcal{Z}_{T,a}$ is a compact subset of $\mathcal{D}([0,T],\Gamma)$, and that d_t is complete over $\mathcal{Z}_{T,a}$.

Let $d_{W,T} : \mathcal{P}\big(\mathcal{E} \times \mathcal{D}([0,T],\mathbb{R})\big) \times \mathcal{P}\big(\mathcal{E} \times \mathcal{D}([0,T],\mathbb{R})\big) \mapsto \mathbb{R}$ be the Wasserstein Distance, i.e

$$d_W(\mu,\nu) = \inf_{\zeta} \big\{ \mathbb{E}^{\zeta}\big[d_{\mathcal{E}}(\theta,\tilde{\theta}) + d_T(x,y) \big] \big\}, \tag{0.2}$$

*Department of Mathematical Sciences. New Jersey Institute of Technology, james.n.maclaurin@njit.edu

where the infimum is over all couplings $\zeta \in \mathcal{P}\big(\mathcal{E} \times \mathcal{D}([0,T],\mathbb{R}) \times \mathcal{E} \times \mathcal{D}([0,T],\mathbb{R})\big)$.

Next define $\mathcal{Y} \subseteq [0,1]^{\Gamma_E \times \Gamma}$ to consist of all $(g_{a,\alpha})_{a\in\Gamma_E,\alpha\in\Gamma}$ such that

$$\sum_{a\in\Gamma_E}\sum_{\alpha\in\Gamma} g_{a,\alpha} = 1. \tag{0.3}$$

Let $B_\epsilon(\theta) \subset \mathcal{E}$ be the open ball about $\theta \in \mathcal{E}$ of radius ϵ.

1. Model Outline. Consider a network of inhomogeneous Poisson Processes on a random adaptive network. Node $j \in I_n$ is assigned a position θ^j_n in a compact Riemannian Manifold $\mathcal{E}$.

HYPOTHESIS 1.1. *There exists a measure $\mu_{\mathcal{E}} \in \mathcal{P}(\mathcal{E})$ such that*

$$\lim_{n\to\infty} n^{-1}\sum_{j\in I_n}\delta_{\theta^j_n} = \mu_{\mathcal{E}}. \tag{1.1}$$

It is also assumed that $\mu_{\mathcal{E}}$ is absolutely continuous with respect to the metric measure on $\mathcal{E}$, with continuous density.

It is assumed that the edges $\{J^{jk}_t\}_{j,k\in I_n\S 1\leq q\leq p}$ take values in a discrete set Γ_E, and that the node variables $\{\sigma^j_t\}_{j\in I_n}$ take values in a discrete state space Γ. We assume that the edge-transitions depend on the states of the end variables, i.e. for $h \ll 1$ and $\alpha \in \Gamma_E$, if $J^{jk}_t = b$ then

$$\mathbb{P}\big(J^{jk}_{t+h} = a \;\big|\; \mathcal{F}_t\big) = h l_{b\mapsto a}\big(\sigma^j_t, \sigma^k_t\big) + O(h^2). \tag{1.2}$$

Define $\hat{\mu}^{n,j}_t \in \mathcal{Y} := \mathcal{P}(\Gamma_E \times \Gamma)$ to represent the local empirical measure containing information about the neighborhood of node j. It is such that, writing, for $\zeta \in \Gamma$ and $a \in \Gamma_E$,

$$\hat{\mu}^{n,j}_t(a,\zeta) = n^{-1}\sum_{k\in I_n}\chi\{J^{jk}_t = a\}\chi\{\sigma^k_t = \zeta\} \tag{1.3}$$

It is assumed that the transitions of the nodes are Poissonian, and such that for all $\alpha \neq \beta$, there exists a Lipschitz function

$$f_{\alpha\mapsto\beta} : \mathcal{Y} \mapsto \mathbb{R}^+ \tag{1.4}$$

such that if $\sigma^j_t = \alpha$ then for $h \ll 1$,

$$\mathbb{P}\big(\sigma^j_{t+h} = \beta \;\big|\; \mathcal{F}_t\big) = h f_{\alpha\mapsto\beta}\big(\hat{\mu}^{n,j}_t\big) + O(h^2). \tag{1.5}$$

HYPOTHESIS 1.2. *We assume that either **(A)** with unit probability, for all $t \geq 0$, $J^{jk}_t = J^{kj}_t$, or that **(B)** the transitions are independent, i.e. for $j \neq k$ and any $T > 0$,*

$$\lim_{h\to 0} h^{-1}\mathbb{P}\big(\text{ For some } t \leq T,\ J^{jk}_{t+h} \neq J^{jk}_t \text{ and } J^{kj}_{t+h} \neq J^{kj}_t\big) = 0. \tag{1.6}$$

The initial conditions $\{\sigma^j_0\}_{j\in I_n}$ and $\{J^{jk}_0\}_{j,k\in I_n}$ are constants (and they can also depend on n, although its neglected from the notation), and it is assumed that the empirical distribution of the initial conditions and spatial locations converges weakly, i.e. it is assumed that

HYPOTHESIS 1.3. *There exists a continuous function $v : \mathcal{E}\times\mathcal{E} \mapsto \mathcal{P}(\Gamma\times\Gamma\times\Gamma_E)$ such that for any continuous function $\mathcal{B} : \mathcal{E}\times\mathcal{E}\times\Gamma\times\Gamma\times\Gamma_E \mapsto \mathbb{R}$,*

$$\sup_{\alpha_1,\alpha_2\in\Gamma}\sup_{a\in\Gamma_E}\lim_{n\to\infty}\bigg|\int_{\mathcal{E}}\int_{\mathcal{E}}\mathcal{B}(x,y,\alpha_1,\alpha_2,a)v_{xy}(\alpha_1,\alpha_2,a)d\mu_{\mathcal{E}}(x)d\mu_{\mathcal{E}}(y) \\ - n^{-2}\sum_{j,k\in I_n}\mathcal{B}(\theta^j_n,\theta^k_n,\alpha_1,\alpha_2,a)\chi\{\sigma^j_0 = \alpha_1, \sigma^k_0 = \alpha_2, J^{jk}_0 = a\}\bigg| = 0. \tag{1.7}$$

1.1. Main Results. We first state our main result.

THEOREM 1.4. *There exists* $\mu \in \mathcal{P}\big(\mathcal{E} \times \mathcal{D}([0,T],\Gamma)\big)$ *(this is precisely specified in Theorem 2.2 below), such that for any* $\epsilon > 0$,

$$\overline{\lim_{n\to\infty}}\, n^{-1} \log \mathbb{P}\big(d_W(\hat{\mu}^n, \mu) \geq \epsilon\big) < 0. \tag{1.8}$$

The limiting probability law μ is (in general) non-autonomous. This makes it difficult to analyze and identify phase transitions and bifurcations (most of the dynamical systems machinery is geared towards autonomous systems). However if we make the additional assumption that the transition rate of the edge J^{jk}_t is independent of σ^j_t (which in general implies that $J^{jk}_t \neq J^{kj}_t$), then we obtain an autonomous expression. This greatly facilitates an analysis of pattern formation and phase transitions.

HYPOTHESIS 1.5. *Suppose that (**A**) of Hypothesis 1.2 holds, and that in addition, the intensity of the connectivity changes only depends on the pre-synaptic activity*

$$l_{b\mapsto a}\big(\sigma^j_t, \sigma^k_t\big) := \tilde{l}_{b\mapsto a}\big(\sigma^k_t\big). \tag{1.9}$$

THEOREM 1.6. *Suppose that Hypothesis 1.5 holds in addition to the other hypotheses. For any continuous function* $\mathcal{H} : \mathcal{E}\times\mathcal{E}\times\Gamma\times\Gamma_E \mapsto \mathbb{R}$ *and any* $T > 0$, $\mathbb{P}$*-almost-surely,*

$$\lim_{n\to\infty} \sup_{\alpha\in\Gamma} \sup_{a\in\Gamma_E} \sup_{t\leq T} \bigg| \int_{\mathcal{E}}\int_{\mathcal{E}} \mathcal{H}(x,y,\alpha,a) p_{xy}(\alpha,a) d\mu_{\mathcal{E}}(x) d\mu_{\mathcal{E}}(y) \\ - n^{-2} \sum_{j,k\in I_n} \mathcal{H}(\theta^j_n, \theta^k_n, \alpha, a)\chi\{\sigma^k_t = \alpha, J^{jk}_t = a\}\bigg| = 0. \tag{1.10}$$

Here $p_{\theta\eta,t}$ *is the unique solution of the following system of PDEs. For* $\theta,\eta\in\mathcal{E}$, *and* $\alpha\in\Gamma$, $a\in\Gamma_E$,

$$p_{\theta\eta,0}(\alpha,a) = \sum_{\beta\in\Gamma} v_{\theta\eta}(\beta,\alpha,a) \tag{1.11}$$

and for $t > 0$,

$$\frac{d}{dt}p_{\theta\eta,t}(\alpha,a) = \sum_{\zeta\in\Gamma:\zeta\neq\beta} \big(f_{\zeta\mapsto\alpha}(G_{\eta,t})p_{\theta\eta,t}(\zeta,a) - f_{\alpha\mapsto\zeta}(G_{\eta,t})p_{\theta\eta,t}(\alpha,a)\big) \\ + \sum_{b\in\Gamma_E\ :b\neq a} \big(\tilde{l}_{b\mapsto a}(\alpha)p_{\theta\eta,t}(\alpha,b) - \tilde{l}_{a\mapsto b}(\alpha)p_{\theta\eta,t}(\alpha,a)\big) \tag{1.12}$$

where $G_{\theta,t} \in \mathcal{P}(\Gamma\times\Gamma_E)$ *is such that for* $\alpha\in\Gamma$, $a\in\Gamma_E$,

$$G_{\theta,t}(\alpha,a) = \int_{\mathcal{E}} p_{\theta\eta,t}(\alpha,a) d\mu_{\mathcal{E}}(\eta) \tag{1.13}$$

Remark 1.7. Theorem 1.6 is equivalent to

$$\lim_{n\to\infty} \sup_{t\leq T} \big\{ d_W\big(\tilde{\mu}^n_t, \nu_t\big)\big\} = 0$$

where

$$\tilde{\mu}^n_t = n^{-2} \sum_{j,k\in I_n} \delta_{\theta^j_n,\theta^k_n,J^{jk}_t,\sigma^k_t} \in \mathcal{P}\big(\mathcal{E}\times\mathcal{E}\times\Gamma_E\times\Gamma\big),$$

and $\nu_t \in \mathcal{P}\big(\mathcal{E}\times\mathcal{E}\times\Gamma_E\times\Gamma\big)$ is such that for measurable $A, B\subseteq\mathcal{E}$, and any $a\in\Gamma_E$ and $\alpha\in\Gamma$,

$$\nu_t(A\times B\times a\times\alpha) = \int_A\int_B p_{\theta\eta,t}(\alpha,a) d\mu_{\mathcal{E}}(\theta) d\mu_{\mathcal{E}}(\eta).$$

We outline the proof of Theorem 1.6.

Proof. A consequence of Hypothesis 1.5 is that the function $\psi_{\theta,t} : \mathcal{D}([0,t],\Gamma) \times \mathcal{P}\big(\mathcal{D}([0,t],\Gamma)\big)$ (defined in (2.7)) is independent of its first argument, and so we can write

$$\tilde{\psi}_{\theta,t}(\mu) := \psi_{\theta,t}(\cdot,\mu).$$

Lemma 2.3 implies that for any $\epsilon > 0$,

$$\overline{\lim_{n\to\infty}} n^{-1} \log \mathbb{P}\big(n^{-1} \sum_{j \in I_n} \sup_{t \le T} d_W\big(\hat{\mu}^{n,j}_t, \tilde{\psi}_{\theta^j_n,t}(\mu)\big) \ge \epsilon\big) < 0, \tag{1.14}$$

and therefore

$$\lim_{n\to\infty} n^{-1} \sum_{j\in I_n} \sup_{t\le T} d_W\big(\hat{\mu}^{n,j}_t, \tilde{\psi}_{\theta^j_n,t}(\mu)\big) = 0.$$

□

2. The Hydrodynamic Limit. We must first specify the limiting probability law for Theorem 1.4. It satisfies a nonlinear implicit equation with delays. We start by specifying how the average connectivity at time t is determined by (i) the initial value of the connectivity and (ii) the trajectories of the afferent spin variables upto time t.

LEMMA 2.1. *For any $c \in \Gamma_E$ any $\theta \in \mathcal{E}$, any $z, y \in \mathcal{D}([0,T],\Gamma)$, there exists a unique $\{j_{\theta,\eta}\}_{\theta,\eta\in\mathcal{E}} \in \mathcal{D}([0,T],\mathcal{P}(\Gamma_E))$, written $j_{\theta\eta} := (j_{\theta\eta,t})_{t\le T}$ such that for all $t \le T$, and all $a \in \Gamma_E$,*

$$j_{\theta\eta,t}(a) = v_{\theta\eta}(z_0, y_0, a) + \sum_{b\in\Gamma_E : b\neq a} \int_0^t \Big\{ j_{\theta\eta,s}(b) l_{b\mapsto a}(z_s, y_s) - j_{\theta\eta,s}(a) l_{a\mapsto b}(z_s,y_s)\Big\} ds \tag{2.1}$$

We write

$$\Psi_{\theta\eta,t} : \mathcal{D}([0,t],\Gamma)^2 \mapsto \mathcal{D}([0,t],\mathcal{Y}) \tag{2.2}$$

$$\Psi_{\theta\eta,t}(z,y) := j_{\theta\eta}. \tag{2.3}$$

Furthermore for each $T > 0$, there exists a constant C_T such that for all $t \le T$,

$$\sup_{\theta,\eta\in\mathcal{E}} \|\Psi_{\theta\eta,t}(z,x) - \Psi_{\theta\eta,t}(\tilde{z},\tilde{x})\| \le C_T\big(\|z-\tilde{z}\|_t + \|x - \tilde{x}\|_t \big). \tag{2.4}$$

Proof. By definition of $\mathcal{D}([0,T],\Gamma)$, both y, z must be piecewise constant, with only a finite number of discontinuities. One then finds a unique solution to the ODE for $j_{\theta,\eta,t}$ along any interval over which both y and z are constant, and iteratively one finds a unique solution upto time T.

To see why (2.4) is true, write $\tilde{j}_{\theta\eta,t}(a) = \Psi_{\theta\eta,t}(\tilde{z},\tilde{x})$. Then one immediately finds that there is a universal constant c such that

$$\sup_{a\in\Gamma} \big| j_{\theta\eta,t}(a) - \tilde{j}_{\theta\eta,t}(a)\big| \le c \int_0^t \sup_{a\in\Gamma} \big| j_{\theta\eta,s}(a) - \tilde{j}_{\theta\eta,s}(a)\big| ds + c\,\|z-\tilde{z}\|_t + c\,\|x-\tilde{x}\|_t \tag{2.5}$$

Gronwall's Inequality now implies

$$\sup_{a\in\Gamma} \big| j_{\theta\eta,t}(a) - \tilde{j}_{\theta\eta,t}(a)\big| \le c\exp(ct)\big(\|z-\tilde{z}\|_t + c\,\|x-\tilde{x}\|_t \big). \tag{2.6}$$

□

For $\theta \in \mathcal{E}$, we next define

$$\psi_{\theta,t} : \mathcal{D}([0,t],\Gamma) \times \mathcal{P}\big(\mathcal{E} \times \mathcal{D}([0,t],\Gamma)\big) \mapsto \mathcal{Y} \tag{2.7}$$

to be such that for $a \in \Gamma_E$ and $\sigma \in \Gamma$,

$$\psi_{\theta,t}(z,\mu)(a,\sigma) = \mathbb{E}^{(\eta,y)\sim\mu}\big[\Psi_{\theta\eta,t}(z,y)(a,\sigma)\chi\{y_t = \sigma\}\big]. \tag{2.8}$$

In the following Theorem we outline an implicit definition of the limiting probability law $\mu_T \in \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,T],\Gamma)\big)$.

THEOREM 2.2. *Let $\{y_{\theta,\alpha\mapsto\beta}(t)\}_{\theta\in\mathcal{E}\S\alpha,\beta\in\Gamma}$ be independent unit intensity Poisson counting processes. For any $T > 0$, there exists a unique set of Γ-valued stochastic processes $\{z_\theta(t)\}_{\theta\in\mathcal{E},t\leq T}$ that satisfy the following properties. Let $\{z_{\theta,0}\}_{\theta\in\mathcal{E}}$ be independent Γ-valued variables, such that*

$$\mathbb{P}\big(z_\theta(0)=\alpha\big)=\sum_{\beta\in\Gamma}\sum_{a\in\Gamma_E}\int_{\mathcal{E}}v_{\theta\eta}(\alpha,\beta,a)d\mu_{\mathcal{E}}(\eta). \tag{2.9}$$

For $t > 0$, define $z_\theta(t) = \alpha \in \Gamma$ precisely when

$$\chi\{z_\theta(0)=\alpha\}+\sum_{\beta\neq\alpha}\Big\{y_{\theta,\beta\mapsto\alpha}\Big(\int_0^t\chi\{z_\theta(s)=\beta\}f_{\beta\mapsto\alpha}\big(\psi_{\theta,s}(z_\theta,\mu_s)\big)ds\Big) \\ - y_{\theta,\alpha\mapsto\beta}\Big(\int_0^t\chi\{z_\theta(s)=\alpha\}f_{\alpha\mapsto\beta}\big(\psi_{\theta,s}(z_\theta,\mu_s)\big)ds\Big)\Big\}=1, \tag{2.10}$$

and in the above $\mu_s \in \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,s],\Gamma)\big)$ is such that for any measurable $A \subset \mathcal{E}$ and any measurable $B \subset \mathcal{D}([0,s],\Gamma)$,

$$\mu_s(A\times B)=\int_A\mathbb{P}\big(z_\theta([0,s])\in B\big)d\mu_{\mathcal{E}}(\theta).$$

Proof. Let $\mathcal{A}_t \subset \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,t],\Gamma)\big)$ consist of all probability measures μ such that, writing μ to be the law of (θ,z), (i) it holds that for any measurable set $B \subset \mathcal{E}$,

$$\mu(\theta\in B)=\mu_{\mathcal{E}}(B).$$

μ must also be such that (ii)

$$\mu\big(\theta\in B,z_0=\alpha\big)=\sum_{\beta\in\Gamma,a\in\Gamma_E}\int_{\mathcal{E}}v_{\theta\eta}(\alpha,\beta,a)d\mu_{\mathcal{E}}(\eta). \tag{2.11}$$

Finally it is required that (iii) we have the following uniform bound on the expected number of transitions: for all $\theta \in \mathcal{E}$,

$$\mathbb{E}^\mu\big[\big|\big\{s\leq t:z(s^-)\neq z(s)\big\}\big|\;\big|\;\theta\big]\leq t\sup_{\alpha,\beta\in\Gamma}\sup_{y\in\mathcal{Y}}f_{\alpha\mapsto\beta}(y) \tag{2.12}$$

where we recall that $f_{\alpha\mapsto\beta}$ is defined in (1.4) and gives the transition intensities.

For some $t > 0$, let $\mu,\tilde{\mu} \in \mathcal{A}_t$ be any two probability measures. We are going to define inhomogeneous counting processes $\{z_\theta(t),\tilde{z}_\theta(t)\}_{\theta\in\mathcal{E}}$. Write

$$g_{\theta,s}=\psi_{\theta,s}(z_\theta,\mu) \tag{2.13}$$

$$\tilde{g}_{\theta,s}=\psi_{\theta,s}(\tilde{z}_\theta,\tilde{\mu}). \tag{2.14}$$

Now define

$$u_{\beta\mapsto\alpha}(t,\theta)=f_{\beta\mapsto\alpha}(g_{\theta,t})\chi\{z_\theta(t)=\beta\} \tag{2.15}$$

$$\tilde{u}_{\beta\mapsto\alpha}(t,\theta)=f_{\beta\mapsto\alpha}(\tilde{g}_{\theta,t})\chi\{\tilde{z}_\theta(t)=\beta\} \tag{2.16}$$

$$\bar{u}_{\beta\mapsto\alpha}(t,\theta)=\inf\big\{u_{\beta\mapsto\alpha}(t,\theta),\tilde{u}_{\beta\mapsto\alpha}(t,\theta)\big\} \tag{2.17}$$

$$\hat{u}_{\beta\mapsto\alpha}(t,\theta)=u_{\beta\mapsto\alpha}(t,\theta)-\bar{u}_{\beta\mapsto\alpha}(t,\theta) \tag{2.18}$$

$$\breve{u}_{\beta\mapsto\alpha}(t,\theta)=u_{\beta\mapsto\alpha}(t,\theta)-\bar{u}_{\beta\mapsto\alpha}(t,\theta). \tag{2.19}$$

For independent unit-intensity Poisson counting processes $\{\bar{Y}_{\beta\mapsto\alpha,\theta}(t), \hat{Y}_{\beta\mapsto\alpha,\theta}(t), \breve{Y}_{\beta\mapsto\alpha,\theta}(t)\}_{\theta\in\mathcal{E}}$, we define

$$\bar{Z}_{\beta\mapsto\alpha,\theta}(t) = \bar{Y}_{\beta\mapsto\alpha,\theta}\Big(\int_0^t \bar{u}_{\beta\mapsto\alpha}(s,\theta)ds\Big) \tag{2.20}$$

$$\hat{Z}_{\beta\mapsto\alpha,\theta}(t) = \hat{Y}_{\beta\mapsto\alpha,\theta}\Big(\int_0^t \hat{u}_{\beta\mapsto\alpha}(s,\theta)ds\Big) \tag{2.21}$$

$$\breve{Z}_{\beta\mapsto\alpha,\theta}(t) = \breve{Y}_{\beta\mapsto\alpha,\theta}\Big(\int_0^t \breve{u}_{\beta\mapsto\alpha}(s,\theta)ds\Big) \tag{2.22}$$

$$Z_{\beta\mapsto\alpha,\theta}(t) = \bar{Z}_{\beta\mapsto\alpha,\theta}(t) + \hat{Z}_{\beta\mapsto\alpha,\theta}(t) \tag{2.23}$$

$$\tilde{Z}_{\beta\mapsto\alpha,\theta}(t) = \bar{Z}_{\beta\mapsto\alpha,\theta}(t) + \breve{Z}_{\beta\mapsto\alpha,\theta}(t). \tag{2.24}$$

Finally it is stipulated that $z_\theta(t) = \beta$ if and only if

$$\chi\{z_\theta(0) = \beta\} + \sum_{\alpha\neq\beta}\big(\bar{Z}_{\alpha\mapsto\beta,\theta}(t) - \bar{Z}_{\beta\mapsto\alpha,\theta}(t)\big) = 1. \tag{2.25}$$

This is well-defined because the LHS of the above equals the number of transitions to β upto time t, minus the number of transitions away, plus one if the initial value is equal to β. We similarly stipulate that $\tilde{z}_\theta(t) = \beta$ if and only if

$$\chi\{z_\theta(0) = \beta\} + \sum_{\alpha\neq\beta}\big(\bar{Z}_{\alpha\mapsto\beta,\theta}(t) - \bar{Z}_{\beta\mapsto\alpha,\theta}(t)\big) = 1. \tag{2.26}$$

Write $\mu_t^{(1)}, \tilde{\mu}_t^{(1)} \in \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,t],\Gamma)\big)$ to be the respective probability laws. That is, for any measurable $A\subseteq\mathcal{E}$, and any measurable $B\subseteq\mathcal{D}([0,t],\Gamma)$,

$$\mu_t^{(1)}(A\times B) = \int_A \mathbb{P}\big(z_\theta\in B\big)d\mu_{\mathcal{E}}(\theta)$$
$$\tilde{\mu}_t^{(1)}(A\times B) = \int_A \mathbb{P}\big(\tilde{z}_\theta\in B\big)d\mu_{\mathcal{E}}(\theta).$$

We next claim that there exists a constant $C>0$ such that

$$\int_{\mathcal{E}}\mathbb{E}\big[d_t(z_\theta,\tilde{z}_\theta)\big] \leq Ctd_W(\mu_t,\tilde{\mu}_t). \tag{2.27}$$

With an aim of proving (2.27), it follows from the definitions that there exists a constant $c>0$ such that for all $t\leq T$,

$$\int_{\mathcal{E}}\mathbb{E}\big[\,\|g_{\theta,t}-\tilde{g}_{\theta,t}\|\,\big]d\theta \leq c\int_{\mathcal{E}}\mathbb{E}\big[\,\|z_\theta-\tilde{z}_\theta\|_t\,\big]d\theta + cd_W(\mu_t,\tilde{\mu}_t). \tag{2.28}$$

Substituting the definitions, and employing Gronwall's Inequality, we then find that there is a constant $C>0$ such that for all $t\leq T$,

$$\begin{aligned}\sup_{\alpha\neq\beta}\int_{\mathcal{E}}\mathbb{E}\big[\hat{Z}_{\beta\mapsto\alpha,\theta}(t)+\breve{Z}_{\beta\mapsto\alpha,\theta}(t)\big]d\theta &\leq C\int_0^t\int_{\mathcal{E}}\mathbb{E}\big[\chi\{z_\theta(s)\neq\tilde{z}_\theta(s)\}\big]d\theta ds + Cd_W(\mu_t,\tilde{\mu}_t)\\ &= C\int_{\mathcal{E}}\mathbb{E}\big[d_t\big(z_\theta,\tilde{z}_\theta\big)\big]d\theta + Cd_W(\mu_t,\tilde{\mu}_t)\end{aligned} \tag{2.29}$$

Notice finally that

$$\int_{\mathcal{E}}\mathbb{E}\big[d_t\big(z_\theta,\tilde{z}_\theta\big)\big]d\theta \leq |\Gamma|^2\sup_{\alpha\neq\beta}\int_{\mathcal{E}}\mathbb{E}\big[\hat{Z}_{\beta\mapsto\alpha,\theta}(t)+\breve{Z}_{\beta\mapsto\alpha,\theta}(t)\big]d\theta \tag{2.30}$$

We can thus conclude that (2.27) holds. (2.27) implies that

$$d_W\big(\mu_t^{(1)},\tilde{\mu}_t^{(1)}\big) \le tC d_W(\mu_t,\tilde{\mu}_t). \tag{2.31}$$

The fixed point theorem then implies that for small enough t, there is a unique μ_t such that $\mu_t^{(1)} = \mu_t$.

We can then iterate this method for larger and larger t, obtaining a unique fixed point upto time T. ∎

2.1. n-dimensional Approximation. We now define an n-dimensional system that approximates the original system, except that the intensity of the flipping of the node-variables $\tilde{\sigma}^j(t)$ is independent of $\tilde{\sigma}^k$ if $k \neq j$, with the same intensity function as the limiting law μ defined in Theorem 2.2. This intermediate approximation will serve as a bridge between the original n-dimensional system and the final limiting law. To this end, let $\{\tilde{\sigma}^j_t\}_{j\in I_n}$ be independent jump-Markov Processes such that (i) $\tilde{\sigma}^j_0 = \sigma^j_0$, (ii) $\tilde{J}^{jk}_0 = J^{jk}_0$, and (iii) for $b \neq a$, if $\tilde{\sigma}^j_t = a$ then

$$\mathbb{P}\big(\tilde{\sigma}^j_{t+h} = \beta \;\big|\; \mathcal{F}_t\big) = h f_{\alpha\mapsto\beta}(\hat{G}^j_t) + O(h^2) \tag{2.32}$$

where

$$\hat{G}^j_t = \psi_{\theta^j_n,t}(\tilde{\sigma}^j,\mu_t) \tag{2.33}$$

and $\mu_t \in \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,t],\Gamma)\big)$ is defined in Theorem 2.2.

The transitions of the connectivities $\{\tilde{J}^{jk}_t\}_{j,k\in I_n}$ are such that for $h \ll 1$ and $\alpha \in \Gamma_E$, if $\tilde{J}^{jk}_t = b$ then

$$\mathbb{P}\big(\tilde{J}^{jk}_{t+h} = a \;\big|\; \mathcal{F}_t\big) = h l_{b\mapsto a}\big(\tilde{\sigma}^j_t,\tilde{\sigma}^k_t\big) + O(h^2). \tag{2.34}$$

Define the associated empirical measure

$$\tilde{\mu}^n_t = n^{-1}\sum_{j\in I_n}\delta_{\theta^j_n,\tilde{\sigma}^j_{[0,t]}} \in \mathcal{P}\big(\mathcal{E}\times\mathcal{D}([0,t],\Gamma)\big). \tag{2.35}$$

Define also, $t \le T$, $\alpha \in \Gamma$ and $a \in \Gamma_E$,

$$\tilde{G}^j_t(\alpha,a) = n^{-1}\sum_{i\in I_n}\chi\big\{\tilde{\sigma}^i_t = \alpha, \tilde{J}^{ji}_t = a\big\} \tag{2.36}$$

Lets first notice that the empirical measure generated by this system converges to the same limiting law.

LEMMA 2.3. *For any $\epsilon > 0$, define the event*

$$\mathcal{U}^n_\epsilon = \Big\{ d_W\big(\tilde{\mu}^n_T,\mu_T\big) \le \epsilon \Big\}. \tag{2.37}$$

Then for any $\epsilon > 0$,

$$\overline{\lim_{n\to\infty}}\, n^{-1}\log\mathbb{P}\big((\mathcal{U}^n_\epsilon)^c\big) < 0. \tag{2.38}$$

Furthermore, for any $\epsilon > 0$,

$$\overline{\lim_{n\to\infty}}\, n^{-1}\log\mathbb{P}\Big(\sup_{j\in I_n}\sup_{t\le T}\sup_{a\in\Gamma_E,\alpha\in\Gamma}\big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| > \epsilon\Big) < 0. \tag{2.39}$$

Proof. The variables $\{\tilde{\sigma}^j\}_{j\in I_n}$ are independent, and the probability law of $\tilde{\sigma}^j$ depends continuously on θ^j_n. It is known that the law of $\tilde{\mu}^n_t$ satisfies a Large Deviation Principle, with rate function $\mathcal{I}:\mathcal{P}\big(\mathcal{D}\times\mathcal{D}([0,T],\Gamma)\big)\mapsto\mathbb{R}$, defined as follows. Suppose first that ν is such that for any measurable $A\subset\mathcal{E}$ and any measurable $B\subset\mathcal{D}([0,T],\Gamma)$, there exists a measurable map $\theta\mapsto\nu_\theta\in\mathcal{P}\big(\mathcal{D}([0,T],\Gamma)\big)$ such that

$$\nu(A\times B)=\int_A \nu_\theta(B)d\mu_{\mathcal{E}}(\theta).$$

In this case, the rate function is such that

$$\mathcal{I}(\nu)=\int_{\mathcal{E}}\mathcal{R}(\nu_\theta||\mu_\theta)d\mu_{\mathcal{E}}(\theta). \tag{2.40}$$

If ν does not admit this decomposition, then $\mathcal{I}(\nu)=\infty$. See for instance [**?**] for a proof.

Since $\mathcal{I}$ has a unique zero at μ, we may conclude that (2.38) holds. It remains to prove (2.39). Define $p^{jk}_t\in\mathcal{Y}$ to be such that $p^{jk}_0(J^{jk}_0)=1$ and $p^{jk}_0(a)=0$ for $a\neq J^{jk}_0$, and for all $a\in\Gamma_E$,

$$p^{jk}_t(a)=p^{jk}_0(a)+\sum_{b\in\Gamma_E:b\neq a}\int_0^t\Big\{p^{jk}_t(b)l_{b\mapsto a}(\tilde{\sigma}^j_s,\tilde{\sigma}^k_s)-p^{jk}_t(a)l_{a\mapsto b}(\tilde{\sigma}^j_s,\tilde{\sigma}^k_s)\Big\}ds. \tag{2.41}$$

Notice also that, since the evolution of $\{\tilde{\sigma}^l\}_{l\in I_n}$ is independent of the evolution of $\{\tilde{J}^{jk}_t\}_{j,k\in I_n}$, for $a\in\Gamma_E$,

$$p^{jk}_t(a)=\mathbb{P}\big(\tilde{J}^{jk}_t=a\;\big|\;\{\tilde{\sigma}^l\}_{l\in I_n}\big). \tag{2.42}$$

Substituting definitions, we find that

$$\tilde{G}^j_t(\alpha,a)-\hat{G}^j_t(\alpha,a)=n^{-1}\sum_{k\in I_n}\chi\{\tilde{\sigma}^k_t=\alpha\}\hat{J}^{jk}_t(a)\text{ where} \tag{2.43}$$

$$\hat{J}^{jk}_t(a)=\chi\{\tilde{J}^{jk}_t=a\}-\mathbb{P}\big(\tilde{J}^{jk}_t=a\;\big|\;\{\tilde{\sigma}^l\}_{l\in I_n}\big). \tag{2.44}$$

Now

$$\mathbb{P}\Big(\sup_{j\in I_n}\sup_{t\leq T}\sup_{a\in\Gamma_E,\alpha\in\Gamma}\big|\tilde{G}^j_t(\alpha,a)-\hat{G}^j_t(\alpha,a)\big|>\epsilon\Big)\leq n^2\sup_{j,p\in I_n}\Big\{\mathbb{P}\Big(\mathcal{V}^{n,j}_\epsilon\big(T(p-1)/n\big)\Big) + |\Gamma||\Gamma_E|\sup_{a\in\Gamma_E,\alpha\in\Gamma}\mathbb{P}\Big(\sup_{(p-1)T/n\leq t\leq pT/n}\big|\tilde{G}^j_t(\alpha,a)-\hat{G}^j_t(\alpha,a)\big|\geq\epsilon/2\Big)\Big\}$$

and we have defined the event, for $j\in I_n$,

$$\mathcal{V}^{n,j}_\epsilon(t)=\Big\{\sup_{\alpha\in\Gamma,a\in\Gamma_E}\big|\tilde{G}^j_t(\alpha,a)-\hat{G}^j_t(\alpha,a)\big|\geq\epsilon/2\Big\}. \tag{2.45}$$

Employing a Chernoff Bound, for a constant $\beta>$,

$$\begin{aligned}\mathbb{P}\big(\mathcal{V}^{n,j}_\epsilon(t)\big)\leq&\sum_{\alpha\in\Gamma,a\in\Gamma_E}\mathbb{E}\Big[\exp\Big(\beta\sum_{k\in I_n}\chi\{\tilde{\sigma}^k_t=\alpha\}\hat{J}^{jk}_t(a)\Big)\\&+\exp\Big(-\beta\sum_{k\in I_n}\chi\{\tilde{\sigma}^k_t=\alpha\}\hat{J}^{jk}_t(a)\Big)\Big]\exp\big(-n\beta\epsilon/2\big)\\ \leq&2|\Gamma||\Gamma_E|\exp\big(C_Tn\beta^2-n\beta\epsilon/2\big)\end{aligned} \tag{2.46}$$

for some constant C_T that is independent of n and β (as long as β is sufficiently small). Substituting $\beta = \epsilon/(4C_T)$, we obtain that

$$\mathbb{P}\big(\mathcal{V}^{n,j}_\epsilon(t)\big) \leq 2|\Gamma||\Gamma_E| \exp\big(-n\epsilon^2/(8C_T)\big) \tag{2.47}$$

Observe that

$$\sup_{(p-1)T/n \leq t \leq pT/n} \big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| \leq n^{-1}\sum_{k\in I_n} \chi\Big\{ \tilde{J}^{jk}_t \neq \tilde{J}^{jk}_{(p-1)T/n} \text{ for some } t \in \big[(p-1)T/n, pT/n\big]\Big\} \tag{2.48}$$

For a constant $\beta > 0$,

$$\mathbb{P}\bigg(\sup_{(p-1)T/n \leq t \leq pT/n} \big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| \geq \epsilon/2\bigg) \leq \mathbb{E}\bigg[\exp\bigg(\beta\sum_{k\in I_n}\chi\Big\{\tilde{J}^{jk}_t \neq \tilde{J}^{jk}_{(p-1)T/n} \text{ for some } t\in\big[(p-1)T/n, pT/n\big]\Big\}\bigg)\bigg]\exp\big(-\beta\epsilon n/2\big)$$

Now the probability that $\tilde{J}^{jk}_t \neq \tilde{J}^{jk}_{(p-1)T/n}$ for some $t \in [(p-1)T/n, pT/n]$ scales as cT/n for a universal constant c. We thus find that

$$\begin{aligned}\mathbb{P}\bigg(\sup_{(p-1)T/n \leq t \leq pT/n} \big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| \geq \epsilon/2\bigg) &\leq \Big(1 + cT/n\big(\exp(\beta)-1\big)\Big)^n \exp\big(-\beta\epsilon n/2\big) \\ &\leq \exp\Big(cT\big(\exp(\beta)-1\big) - \beta\epsilon n/2\Big).\end{aligned}$$

Choosing $\beta = \log n$, we find that for large enough n,

$$\sup_{p,j\in I_n}\mathbb{P}\bigg(\sup_{(p-1)T/n \leq t \leq pT/n} \big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| \geq \epsilon/2\bigg) \leq \exp\big(-n\big)$$

Combining the above estimates, we find that

$$\overline{\lim_{n\to\infty}}\, n^{-1}\log\mathbb{P}\bigg(\sup_{j\in I_n}\sup_{t\leq T}\sup_{a\in\Gamma_E,\alpha\in\Gamma}\big|\tilde{G}^j_t(\alpha,a) - \hat{G}^j_t(\alpha,a)\big| > \epsilon\bigg) < 0,$$

as required. ∎

2.2. Coupling of Systems. Next, we outline a coupling of the n-dimensional systems in the same space (this roughly means that the probability laws of $\{\sigma^j_t\}_{j\in I_n}$ and $\{\tilde{\sigma}^j_t\}_{j\in I_n}$ remain the same, but the system is constructed to maximize the correlations as much as possible). To do this, we will employ the time-rescaling formalism that has been extensively employed by Kurtz [9] and Anderson [2] (amongst others).

To this end, let

$$\big\{\grave{Y}^j_{\alpha\mapsto\beta}(\cdot), \bar{Y}^j_{\alpha\mapsto\beta}(\cdot), \check{Y}^j_{\alpha\mapsto\beta}(\cdot)\big\}_{\alpha,\beta\in\Gamma ⨾ \alpha\neq\beta ⨾ j\in I_n} \subset \mathcal{D}\big([0,\infty),\mathbb{Z}^+\big) \tag{2.49}$$

be independent unit intensity Poisson Processes. For the edge dynamics, we analogously let

$$\big\{\grave{Y}^{jk}_{a\mapsto b}(\cdot), \bar{Y}^{jk}_{a\mapsto b}(\cdot), \check{Y}^{jk}_{a\mapsto b}(\cdot)\big\}_{a,b\in\Gamma_E ⨾ a\neq b ⨾ j,k\in I_n} \tag{2.50}$$

be another set of mutually independent unit intensity Poisson Processes. We are going to define the stochastic processes $\{\sigma^j_t, \tilde{\sigma}^j_t, \hat{G}^j_{\zeta,t}\}_{j\in I_n,\zeta\in\mathcal{E}}$ to be functions of the above processes. To this end, we first note that the initial conditions are as previously specified, i.e. $\tilde{\sigma}^j_0 = \sigma^j_0$, $\tilde{J}^{jk}_0 = J^{jk}_0$. As previously, we define

$$\hat{G}^j_t = \psi_{\theta^j_n, t}(\tilde{\sigma}^j, \mu_t) \tag{2.51}$$

We next define $\big\{Z^j_{\alpha\mapsto\beta}(t), \tilde{Z}^j_{\alpha\mapsto\beta}(t), \grave{Z}^j_{\alpha\mapsto\beta}(t), \breve{Z}^j_{\alpha\mapsto\beta}(t)\big\}_{j\in I_n}$ to be counting processes, and such that

$$\begin{aligned}
Z^j_{\alpha\mapsto\beta}(t) =& \grave{Z}^j_{\alpha\mapsto\beta}(t) + \bar{Y}^j_{\alpha\mapsto\beta}\bigg(\int_0^t \bar{f}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.52)\\
\tilde{Z}^j_{\alpha\mapsto\beta}(t) =& \breve{Z}^j_{\alpha\mapsto\beta}(t) + \bar{Y}^j_{\alpha\mapsto\beta}\bigg(\int_0^t \bar{f}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.53)\\
\grave{Z}^j_{\alpha\mapsto\beta}(t) =& \grave{Y}^j_{\alpha\mapsto\beta}\bigg(\int_0^t \grave{f}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.54)\\
\breve{Z}^j_{\alpha\mapsto\beta}(t) =& \breve{Y}^j_{\alpha\mapsto\beta}\bigg(\int_0^t \breve{f}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.55)\\
\bar{f}^j_{\alpha\mapsto\beta}(t) =& \inf\big\{f_{\alpha\mapsto\beta}(G^j_t)\chi\{\sigma^j_t = \alpha\}, f_{\alpha\mapsto\beta}(\hat{G}^j_t)\chi\{\tilde{\sigma}^j_t = \alpha\}\big\} && (2.56)\\
\grave{f}^j_{\alpha\mapsto\beta}(t) =& f_{\alpha\mapsto\beta}(G^j_t)\chi\{\sigma^j_t = \alpha\} - \bar{f}^j_{\alpha\mapsto\beta}(t) && (2.57)\\
\breve{f}^j_{\alpha\mapsto\beta}(t) =& f_{\alpha\mapsto\beta}(\hat{G}^j_t)\chi\{\tilde{\sigma}^j_t = \alpha\} - \bar{f}^j_{\alpha\mapsto\beta}(t). && (2.58)
\end{aligned}$$

The edge variables are coupled in an analogous manner to the node variables.

$$\begin{aligned}
Z^{jk}_{a\mapsto b}(t) =& \grave{Z}^{jk}_{a\mapsto b}(t) + \bar{Y}^{jk}_{a\mapsto b}\bigg(\int_0^t \bar{l}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.59)\\
\tilde{Z}^{jk}_{a\mapsto b}(t) =& \breve{Z}^{jk}_{a\mapsto b}(t) + \bar{Y}^{jk}_{a\mapsto b}\bigg(\int_0^t \bar{l}^j_{\alpha\mapsto\beta}(s)ds\bigg) && (2.60)\\
\grave{Z}^{jk}_{a\mapsto b}(t) =& \grave{Y}^{jk}_{a\mapsto b}\bigg(\int_0^t \grave{l}^{jk}_{a\mapsto b}(s)ds\bigg) && (2.61)\\
\breve{Z}^{jk}_{a\mapsto b}(t) =& \breve{Y}^{jk}_{a\mapsto b}\bigg(\int_0^t \breve{l}^{jk}_{a\mapsto b}(s)ds\bigg) && (2.62)\\
\bar{l}^{jk}_{a\mapsto b}(t) =& \inf\big\{l_{a\mapsto b}(\sigma^j_t, \sigma^k_t)\chi\{J^{jk}_t = a\}, l_{a\mapsto b}(\tilde{\sigma}^j_t, \tilde{\sigma}^k_t)\chi\{\tilde{J}^{jk}_t = a\}\big\} && (2.63)\\
\grave{l}^{jk}_{a\mapsto b}(t) =& l_{a\mapsto b}(\sigma^j_t, \sigma^k_t)\chi\{J^{jk}_t = a\} - \bar{l}^{jk}_{a\mapsto b}(t) && (2.64)\\
\breve{l}^{jk}_{a\mapsto b}(t) =& l_{a\mapsto b}(\tilde{\sigma}^j_t, \tilde{\sigma}^k_t)\chi\{\tilde{J}^{jk}_t = a\} - \bar{l}^{jk}_{a\mapsto b}(t). && (2.65)
\end{aligned}$$

We then stipulate that, for any $\alpha \in \Gamma$,

$$\begin{aligned}
& \sigma^j_t = \alpha \text{ if and only if} && (2.66)\\
\gamma^j_t(\alpha) :=& \chi\{\sigma^j_0 = \alpha\} + \sum_{\beta\neq\alpha}\big(Z^j_{\beta\mapsto\alpha}(t) - Z^j_{\alpha\mapsto\beta}(t)\big) = 1 && (2.67)
\end{aligned}$$

Note that only one of $\{\gamma^j_t(\alpha)\}_{\alpha\in\Gamma}$ can equal one, and the rest must be zero. This is because $\gamma^j_t(\alpha)$ counts the net number of transitions to the state α minus the number of transitions away, plus 1 if σ^j_0 is α. We similarly stipulate that

$$\begin{aligned}
& \tilde{\sigma}^j_t = \alpha \text{ if and only if} && (2.68)\\
& \chi\{\sigma^j_0 = \alpha\} + \sum_{\beta\neq\alpha}\big(\tilde{Z}^j_{\beta\mapsto\alpha}(t) - \tilde{Z}^j_{\alpha\mapsto\beta}(t)\big) = 1. && (2.69)
\end{aligned}$$

For the edge variables, we stipulate that, for any $a \in \Gamma_E$

$$J^{jk}_{a\mapsto b}(t) = a \text{ in and only if} \tag{2.70}$$

$$\chi\{J^{jk}_0 = a\} + \sum_{b\neq a}\big(Z^{jk}_{b\mapsto a}(t) - Z^{jk}_{a\mapsto b}(t)\big) = 1 \tag{2.71}$$

and we stipulate that

$$\tilde{J}^{jk}_{a\mapsto b}(t) = a \text{ if and only if} \tag{2.72}$$

$$\chi\{J^{jk}_0 = a\} + \sum_{b\neq a}\big(\tilde{Z}^{jk}_{b\mapsto a}(t) - \tilde{Z}^{jk}_{a\mapsto b}(t)\big) = 1. \tag{2.73}$$

LEMMA 2.4. *The above system is well-defined and consistent with the earlier definitions.*

Proof. The fact that the variables $\big\{Z^j_{\alpha\mapsto\beta}(t), \tilde{Z}^j_{\alpha\mapsto\beta}(t), \grave{Z}^j_{\alpha\mapsto\beta}(t), \breve{Z}^j_{\alpha\mapsto\beta}(t), \sigma^j_t, \tilde{\sigma}^j_t\big\}_{j\in I_n}$ are well-defined follows from the fact that the counting processes (2.49) - (2.50) are piecewise-constant, with only a finite number of jumps over a finite time interval. For the time-intervals between jumps, there is thus existence and uniqueness due to the Picard-Lindelof Theorem for ODEs. We note also that these stochastic variables are adapted to the same filtration. ∎

We need to control the distance between the two systems. To this end, for $t \leq T$, define

$$\delta^n_t = n^{-1} z^n_t \text{ where} \tag{2.74}$$

$$z^n_t = \sum_{j\in I_n}\sum_{\alpha,\beta\in\Gamma:\alpha\neq\beta}\big(\grave{Z}^j_{\alpha\mapsto\beta}(t) + \breve{Z}^j_{\alpha\mapsto\beta}(t)\big) \tag{2.75}$$

$$u^{n,j}_t = \sum_{k\in I_n}\sum_{a,b\in\Gamma_E:a\neq b}\big(\grave{Z}^{jk}_{a\mapsto b}(t) + \breve{Z}^{jk}_{a\mapsto b}(t)\big) \tag{2.76}$$

$$\varphi^n_t = n^{-2} u^n_t \text{ where} \tag{2.77}$$

$$u^n_t = \sum_{j,k\in I_n}\sum_{a,b\in\Gamma_E:a\neq b}\big(\grave{Z}^{jk}_{a\mapsto b}(t) + \breve{Z}^{jk}_{a\mapsto b}(t)\big) \tag{2.78}$$

$$\eta^n_t = n^{-1}\sum_{j\in I_n}\|\tilde{G}^j_t - \hat{G}^j_t\| \tag{2.79}$$

The main result that we must prove in this section is the following Lemma.

LEMMA 2.5. *For any $\epsilon > 0$, there exists $\tilde{\epsilon} > 0$ such that*

$$\overline{\lim_{n\to\infty}} n^{-1}\log\mathbb{P}\big(\delta^n_T \geq \epsilon \;\;,\;\; \sup_{t\leq T}\eta^n_t \leq \tilde{\epsilon}\big) < 0. \tag{2.80}$$

We next notice how Lemma 2.5 implies the veracity of Theorem 1.4.

Proof. Thanks to the triangle inequality

$$d_W\big(\hat{\mu}^n_T, \mu_T\big) \leq d_W\big(\tilde{\mu}^n_T, \mu_T\big) + d_W\big(\tilde{\mu}^n_T, \hat{\mu}^n_T\big).$$

Hence

$$\overline{\lim_{n\to\infty}} n^{-1}\log\mathbb{P}\big(d_W\big(\hat{\mu}^n_T, \mu_T\big) \geq \epsilon, \sup_{t\leq T}\eta^n_t \leq \tilde{\epsilon}\big) \leq \tag{2.81}$$
$$\max\Big\{\overline{\lim_{n\to\infty}} n^{-1}\log\mathbb{P}\Big(d_W\big(\tilde{\mu}^n_T, \mu_T\big) \geq \epsilon/2, \sup_{t\leq T}\eta^n_t \leq \tilde{\epsilon}\Big),$$
$$\overline{\lim_{n\to\infty}} n^{-1}\log\mathbb{P}\Big(d_W\big(\hat{\mu}^n_T, \tilde{\mu}^n_T\big) \geq \epsilon/2, \sup_{t\leq T}\eta^n_t \leq \tilde{\epsilon}\Big)\Big\}$$

Thanks to Lemma 2.3, the first term on the RHS is negative. Lemma 2.5 implies that the second term on the RHS of (2.81) is negative. We therefore find that

$$\overline{\lim_{n\to\infty}} n^{-1}\log \mathbb{P}\big(d_W\left(\hat{\mu}^n_T,\mu_T\right)\geq \epsilon, \sup_{t\leq T}\eta^n_t \leq \tilde{\epsilon}\big) < 0.$$

∎

The rest of this paper is dedicated to proving Lemma 2.5. To this end, we must first prove that the intensities of the 'remainder' Poisson Processes can be uniformly bounded in terms of δ^n_t and $\varphi^{n,j}_t$.

Lemma 2.6. *There exists a constant c_T (the constant is independent of n and j) such that for all $t\leq T$, all $j\in I_n$,*

$$n^{-1}\sum_{k\in I_n}\sum_{\alpha,\beta\in\Gamma\,:\,\alpha\neq\beta}\big(\grave{f}^k_{\alpha\mapsto\beta}(t)+\breve{f}^k_{\alpha\mapsto\beta}(t)\big)\leq c_T\big(\delta^n_t+\varphi^n_t+\eta^n_t\big) \tag{2.82}$$

$$n^{-2}\sum_{j,k\in I_n}\sum_{a,b\in\Gamma_E\,:\,a\neq b}\big(\grave{l}^{jk}_{a\mapsto b}(t)+\breve{l}^{jk}_{a\mapsto b}(t)\big)\leq c_T\big(\delta^n_t+\varphi^n_t\big) \tag{2.83}$$

Proof. Since the function $f_{\alpha\mapsto\beta}$ is Lipschitz and bounded, there is a constant $c>0$ such that

$$n^{-1}\sum_{k\in I_n}\sum_{\alpha,\beta\in\Gamma\,:\,\alpha\neq\beta}\big(\grave{f}^k_{\alpha\mapsto\beta}(t)+\breve{f}^k_{\alpha\mapsto\beta}(t)\big)\leq n^{-1}c\sum_{k\in I_n}\big(\|G^k_t-\hat{G}^k_t\|+\chi\{\sigma^k_t\neq\tilde{\sigma}^k_t\}\big) \tag{2.84}$$

Furthermore

$$n^{-1}\sum_{k\in I_n}\chi\{\sigma^k_t\neq\tilde{\sigma}^k_t\}\leq|\Gamma|\delta^n_t$$

Furthermore, thanks to the triangle inequality,

$$\begin{aligned}\|G^k_t-\hat{G}^k_t\| \leq&\|\tilde{G}^k_t-\hat{G}^k_t\|+\|G^k_t-\tilde{G}^k_t\|\\ \leq&\|\tilde{G}^k_t-\hat{G}^k_t\|+\mathrm{Const}\times\mathrm{n}^{-1}\sum_{\mathrm{q}\in\mathrm{I_n}}\chi\{\sigma^{\mathrm{q}}_{\mathrm{t}}\neq\tilde{\sigma}^{\mathrm{q}}_{\mathrm{t}}\}+\mathrm{Const}\times\mathrm{n}^{-1}\mathrm{u}^{\mathrm{n,k}}_{\mathrm{t}}\\ \leq&\|\tilde{G}^k_t-\hat{G}^k_t\|+\mathrm{Const}\times\mathrm{n}^{-1}\mathrm{u}^{\mathrm{n,k}}_{\mathrm{t}}+\mathrm{Const}\times\delta^{\mathrm{n}}_{\mathrm{t}}.\end{aligned} \tag{2.85}$$

Combining these bounds, we obtain (2.82). The proof of (2.83) is analogous. ∎

For $\tilde{\epsilon}>0$, define the event

$$\mathcal{W}_{n,\tilde{\epsilon}}=\Big\{\sup_{t\leq T}\eta^n_t\leq\tilde{\epsilon}\Big\} \tag{2.86}$$

Now suppose that $y(t),q(t)$ are independent unit intensity counting proceses, and for a constant $c>0$, define the counting processes to be such that

$$\tilde{z}_n(t)=y\bigg(ncT\tilde{\epsilon}+c\int_0^t\big\{\tilde{z}_n(s)+n^{-1}\tilde{u}_n(s)\big\}ds\bigg) \tag{2.87}$$

$$\tilde{u}_n(t)=q\bigg(c\int_0^t n\tilde{z}_n(s)ds+\tilde{u}_n(s)ds\bigg) \tag{2.88}$$

Lemma 2.7. *As long as the constant $c>0$ is sufficiently large, it holds that*

$$\mathbb{P}\bigg(\max\{\delta^n_T,\phi^n_T\}\geq\epsilon\;,\;\mathcal{W}_{n,\tilde{\epsilon}}\bigg)\leq\mathbb{P}\bigg(\max\{n^{-1}\tilde{z}_n(T),n^{-2}\tilde{u}_n(T)\}\geq\epsilon\bigg) \tag{2.89}$$

Proof. We define $c = c_T$, where c_T is specified in Lemma 2.6. Now $z_n(t)$ and $u_n(t)$ are inhomogeneous counting processes (i.e. they increase by increments of 1, with random intensities). Substituting the upper bounds for the intensities that are obtained in Lemma 2.6, we obtain (2.87) - (2.88). This implies the Lemma. ∎

For a large positive integer m, let $t_i^{(m)} = iT/m$ and for any $t \in [0,T]$, $t^{(m)} := \sup\{s \le t \ : s = t_i^{(m)} \text{ for some } i \le m\}$. Write

$$p_n(t) = \max\{n^{-1}\tilde{z}_n(t), n^{-2}\tilde{u}_n(t)\} \tag{2.90}$$

For a constant $a > 0$, let $\tilde{\epsilon}_m > 0$ be such that (i)

$$\tilde{\epsilon}_m > T\tilde{\epsilon}/m + T\epsilon \exp\big(aT\big)/m$$

and (ii) $\lim_{m\to\infty} \tilde{\epsilon}_m = 0$. Define the stopping times

$$\begin{aligned} \tau_m = &\inf\big\{t \le T \ : \ \tilde{z}_n(t) - \tilde{z}_n\big(t^{(m)}\big) \ge n\tilde{\epsilon}_m \text{ and } \tilde{z}_n(t) \le \tilde{\epsilon}_m + \epsilon\exp(aT) \\ & \quad \text{or } \tilde{u}_n(t) - \tilde{u}_n\big(t^{(m)}\big) \ge n^2\tilde{\epsilon}_m \text{ and } \tilde{u}_n(t) \le \tilde{\epsilon}_m + \epsilon\exp(aT)\big\} \end{aligned} \tag{2.91}$$

$$\tilde{\tau}_m = \inf\big\{t \le T \ : \ t = t_a^{(m)} \text{ and } p_n(t) \ge \epsilon\exp(at)\big\} \tag{2.92}$$

Lemma 2.8. *For all $m \ge 2$,*

$$\begin{aligned} \varlimsup_{n\to\infty} n^{-1}\log\mathbb{P}\Bigg(&\sup_{0\le b\le m-1}\ \sup_{t\in[t_b^{(m)}, t_{b+1}^{(m)}]} \big\{\tilde{z}_n\big(t\big) - \tilde{z}_n\big(t_b^{(m)}\big)\big\} \ge n\tilde{\epsilon}_m \ \textit{or} \\ &\sup_{0\le b\le m-1}\ \sup_{t\in[t_b^{(m)}, t_{b+1}^{(m)}]} \big\{\tilde{u}_n\big(t\big) - \tilde{u}_n\big(t_b^{(m)}\big)\big\} \ge n^2\tilde{\epsilon}_m\ , \quad \textit{and } t \le \tau\wedge\tilde{\tau}_m\Bigg) < 0. \end{aligned} \tag{2.93}$$

Proof. For all times $t \le \tau\wedge\tilde{\tau}_m$, $\tilde{z}_n\big(t\big) - \tilde{z}_n\big(t_b^{(m)}\big)$ is Poissonian, with intensity upperbounded by

$$n\big(\tilde{\epsilon}_m + \epsilon\exp(aT)\big).$$

Similarly, $\tilde{u}_n\big(t\big) - \tilde{u}_n\big(t_b^{(m)}\big)$ is Poissonian, with intensity upperbounded by

$$n^2\big(\tilde{\epsilon}_m + \epsilon\exp(aT)\big).$$

Since $\tilde{\epsilon}_m > \frac{T}{m}\big(\tilde{\epsilon}_m + \epsilon\exp(aT)\big)$, the result is now a standard result from Poisson Processes. ∎

Corollary 2.9.

$$\varlimsup_{n\to\infty} n^{-1}\log\mathbb{P}\big(\tau_m < T\big) < 0.$$

Lemma 2.10.

$$\varlimsup_{n\to\infty} n^{-1}\log\mathbb{P}\big(\tilde{\tau}_m < T \ \textit{and } \tau_m \ge T\big) < 0.$$

Proof. For all times $t \le \tau_m$, $\tilde{z}_n(t)$ is Poissonian, and is such that

$$\tilde{z}_n(t) \le y\bigg(2cn\big(\tilde{\epsilon}_m + \epsilon a^{-1}\exp(at)\big)\bigg).$$

Similarly, for all times $t \le \tau_m$, $\tilde{u}_n(t)$ is Poissonian, with intensity upperbounded by

$$\tilde{u}_n(t) \le q\bigg(2cn^2\big(\tilde{\epsilon}_m + \epsilon a^{-1}\exp(at)\big)\bigg).$$

Standard Poisson Identities dictate that, as long as (i) $2c\tilde{\epsilon}_m \leq \epsilon \exp(aT)$ and (ii) a is such that $2c/a < 1/2$, it must hold that

$$\overline{\lim_{n\to\infty}} n^{-1} \log \mathbb{P}\Big(\text{ For some } t = t_b^{(m)},\ y\Big(2cn\big(\tilde{\epsilon}_m + \epsilon a^{-1}\exp(at)\big)\Big) \geq \epsilon n \exp(at)\Big) < 0. \tag{2.94}$$

Similarly,

$$\overline{\lim_{n\to\infty}} n^{-1} \log \mathbb{P}\Big(\text{ For some } t = t_b^{(m)},\ q\Big(2cn^2\big(\tilde{\epsilon}_m + \epsilon a^{-1}\exp(at)\big)\Big) \geq \epsilon n^2 \exp(at)\Big) < 0. \tag{2.95}$$ □

Observe that the intensity of the process over the time interval

LEMMA 2.11. *For any $\epsilon > 0$,*

$$\overline{\lim_{n\to\infty}} n^{-1} \log \mathbb{P}\Big((\delta^n_T + \phi^n_T) \geq \epsilon \Big) < 0. \tag{2.96}$$

Proof. For the constant $c_T > 0$ of Lemma, define the counting process $z^j_t \in \mathcal{D}([0,T],\mathbb{Z})$ to be such that $z_0 = 1$, and for independent unit-intensity counting processes $\{y^j(\cdot)\}_{j\in I_n}$,

$$z^j_t = y\Big(2c_T \int_0^t z^j_s ds\Big). \tag{2.97}$$

We next claim that

$$\mathbb{P}\Big((\delta^n_T + \phi^n_T) \geq \epsilon, \eta^n_T \leq \tilde{\epsilon} \Big) \leq \mathbb{P}\Big(\sum_{j=1}^{\lfloor Tc_T\tilde{\epsilon}n \rfloor} z^j_t \geq n\epsilon \Big) \tag{2.98}$$ □

Let $m \gg 1$ be a positive integer, and write $t_i^{(m)} = iT/m$. For some $\epsilon \ll 1$, define the stopping time

$$\tau^n_\epsilon = \inf\big\{ t \leq T \ :\ \delta^n_t = \epsilon \text{ or } \varphi^n_t = \epsilon \big\} \tag{2.99}$$

LEMMA 2.12. *For any $\epsilon < 1$, writing $q^n_t = n^{-1}(z^n_t + u^n_t)$, it holds that*

$$\overline{\lim_{n\to\infty}} n^{-1} \log \mathbb{P}\big(\sup_{0\leq a\leq m-1} (q^n_{t^{(m)}_{a+1}} - q^n_{t^{(m)}_a}) \geq 2\epsilon/m \big) < 0. \tag{2.100}$$

Define the stopping time, for a constant $c > 0$

$$\tilde{\tau}^n = \inf\big\{ t \leq \tau^n_\epsilon \ :\ \big\} \tag{2.101}$$

REFERENCES

[1] Zoé Agathe-Nerine. Multivariate hawkes processes on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 152:86–148, 10 2022.
[2] David F Anderson and Thomas G Kurtz. *Stochastic Analysis of Biochemical Systems*. Springer, 2015.
[3] D Avitabile and J Maclaurin. Large deviations of hawkes processes on structured sparse random graphs. 2024.
[4] Riddhipratim Basu and Allan Sly. Evolving voter model on dense random graphs. *Annals of Applied Probability*, 27:1235–1288, 4 2017.
[5] Jason Bramburger and Matt Holzer. Pattern formation in random networks using graphons. *SIAM Journal on Mathematical Analysis*, 55:2150–2185, 2023.
[6] Jason J. Bramburger, Matt Holzer, and Jackson Williams. Persistence of steady-states for dynamical systems on large networks. 2 2024.
[7] J. Chevallier, A. Duarte, E. Löcherbach, and G. Ost. Mean field limits for nonlinear spatially extended hawkes processes with exponential memory kernels. *Stochastic Processes and their Applications*, 129:1–27, 2019.

[8] Fabio Coppini, Eric Luçon, and Christophe Poquet. Central limit theorems for global and local empirical measures of diffusions on erdos-renyi graphs. *Electronic Journal of Probability*, 28, 2023.
[9] Stewart Ethier and Thomas Kurtz. *Markov Processes. Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics, 1986.
[10] Marios Antonios Gkogkas, Christian Kuehn, and Chuang Xu. Mean field limits of co-evolutionary heterogeneous networks. *Arxiv Preprint*, 2023.
[11] Thilo Gross and H Sayama. *Adaptive Networks*. Springer, 2009.
[12] Christian Kuehn. Time-scale and noise optimality in self-organized critical adaptive networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85, 2 2012.
[13] Laszlo Lovasz. *Large Networks and Graph Limits*. 2012.
[14] Eric Lucon. Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 130:6783–6842, 11 2020.
[15] James Maclaurin. Large deviations of a network of interacting particles with sparse random connections. *Submittied to SIAM Dynamical Systems*, 2024.
[16] Remco van der Hofstad. *Random Graphs and Complex Networks. Volume 2*. Cambridge, 2024.