



PROPAGATION OF CHAOS AND LARGE DEVIATIONS IN MEAN-FIELD MODELS WITH JUMPS ON BLOCK-STRUCTURED NETWORKS

DONALD A. DAWSON,* **
AHMED SID-ALI,* *** AND
YIQIANG Q. ZHAO,* **** *Carleton University*

Abstract

A system of interacting multi-class finite-state jump processes is analyzed. The model under consideration consists of a block-structured network with dynamically changing multi-color nodes. The interactions are local and described through local empirical measures. Two levels of heterogeneity are considered: between and within the blocks where the nodes are labeled into two types. The central nodes are those connected only to nodes from the same block, whereas the peripheral nodes are connected to both nodes from the same block and nodes from other blocks. Limits of such systems as the number of nodes tends to infinity are investigated. In particular, under specific regularity conditions, propagation of chaos and the law of large numbers are established in a multi-population setting. Moreover, it is shown that, as the number of nodes goes to infinity, the behavior of the system can be represented by the solution of a McKean–Vlasov system. Then, we prove large deviations principles for the vectors of empirical measures and the empirical processes, which extends the classical results of Dawson and Gärtner (*Stochastics* **20**, 1987) and Léonard (*Ann. Inst. H. Poincaré Prob. Statist.* **31**, 1995).

Keywords: Interacting particle systems; inhomogeneous graphs; McKean–Vlasov equation; multi-class populations; jump processes

2020 Mathematics Subject Classification: Primary 60K35; 60J27; 60F10
Secondary 05C82; 60J74

1. Introduction

Since McKean’s seminal paper [51], mean-field theory has been widely used to study large stochastic interacting particle systems arising in various domains such as statistical physics [21, 37, 51, 52], biological systems [23, 53], communication networks [8, 39, 41, 42], mathematical finance [38, 47], etc. This theory, first initiated in connection with the mathematical foundation of the Boltzmann equation, aims for a mathematically rigorous treatment of the time evolution of stochastic systems with weak long-range interaction where the interaction between the particles is realized via the *empirical measure* of the particle configuration. For such scenarios, it is then natural to investigate the behavior of the *empirical process* instead

Received 23 March 2021; revision received 6 February 2023.

* Postal address: School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.

** Email address: ddawson@math.carleton.ca

*** Email address: ahmedsidali@cunet.carleton.ca

**** Email address: zhao@math.carleton.ca

© The Author(s), 2023. Published by Cambridge University Press on behalf of Applied Probability Trust.

of considering the particle configuration itself. In particular, one is interested in investigating laws of large numbers, limit theorems, and large deviations for the empirical process in the limit as the number of particles tends to infinity. Another concept that plays an important role is *chaos propagation*, first introduced by Kac as part of kinetic theory [45], and then widely studied in the literature since; see, e.g., [37] and [61] for detailed developments on the subject.

Classically, the systems studied are homogeneous with complete interaction graphs; that is, the particles are exchangeable and each particle interacts with all the others. In such a setting, the big picture is well understood and various asymptotic results have been established for a variety of models. One can consult [22], [37], or [61] for an overview. However, though such assumptions are reasonable in statistical physics and accurately describe a variety of phenomena, this may no longer be the case when considering other applications. Researchers have therefore studied many new interacting particle systems where the homogeneity or complete interaction assumptions are not tenable; this is the direction in which this paper proceeds.

Systems in which particles carry intrinsic distinguishing features lead naturally to heterogeneous models. Thus, one cannot presume the particles to be identically distributed. Instead, one relies on additional conditions to establish limiting results. For instance, in [34, 35, 38], models for the activities of heterogeneous economic agents were proposed and laws of large numbers were proved under some regularity conditions. In [18], the authors investigated systems of interacting stochastic differential equations with two kinds of heterogeneity: one originating from different weights of the linkages, and the other concerning their asymptotic relevance when the system becomes large. The authors then introduced a partial mean-field system by averaging only over the particles with weak interactions and proved a law of large numbers together with a large deviations principle.

A particular instance of heterogeneity, with which our current work is in line, is the multi-population paradigm, also known as multi-types or multi-species. Here, the particles of the system are divided into groups, within which they are homogeneous or partially homogeneous. The interest in these structures is motivated by their ubiquity in different fields (see references below). We thus propose in this paper to take a step forward in the understanding of their behavior by studying the large-scale asymptotics of interacting particle systems with jumps on block-structured networks. Specifically, we will set up a model for block-structured networks with dynamically changing multi-color nodes. The evolution of node colors is described by a sequence of finite-state pure-jump processes interacting through local empirical measures describing the neighborhood of each node. The nodes of the network are divided into a finite number of blocks. In addition, the nodes within each block are divided into two subgroups: central and peripheral nodes. The central nodes are those connected only to nodes from the same block, whereas the peripheral nodes interact with both particles from the same block and some particles from other blocks. Thus, our model describes two levels of heterogeneity: between blocks and within blocks.

Block-structured networks are ubiquitous in various interacting particle systems composed of different communities, where a given community consists of a group of agents densely connected to each other but sparsely connected to the other dense groups of the population. For instance, a community in a social network might refer to a circle of friends, a community on the World Wide Web might include a group of pages on closely related topics, and a community in a cellular or genetic network might be related to a functional module. The study of the grouping patterns of communities, together with their detection, is an active field of research among physicists and applied mathematicians, and the study of what has become known as

community structure is now one of the most prominent areas of network science. The reader can consult [59] and the references therein for an overview of the subject.

Our idea builds on the results in several existing works. The authors of [19, 20] studied a bi-populated Curie–Weiss model and established, via a large deviations approach, the propagation of chaos and the asymptotic dynamics of the pair of group magnetizations in the infinite-volume limit. Laws of large numbers and a central limit theorem were proved in [46] for an extension of this model to the case of heterogeneous coupling within and between groups. A related paper [50] studied high-temperature fluctuations for a block spin Ising model and established a central limit theorem. A variant of this model was analyzed in [48], where the vertices are divided into a finite number of blocks and pair interactions are given according to their blocks. The authors proved a large deviations principle and a central limit theorem. In the same spirit, limiting results were established in [56] for a system of reflected diffusions segregated into two groups of blue and red particles and subject to a reflection condition. These results were extended in [57] to the case with drift coefficients not of average form. Other recent related works are [54], which studied the two-community noisy Kuramoto model, and [4], which studied opinion dynamics in a model with Lotka–Volterra-type interactions. Among other instances of the multi-population paradigm, we mention in particular the works [14, 17], which considered mean-field game models with a single major player and statistically identical minor players. Propagation of chaos was proved for the minor players conditioned on the major player.

Another closely related model was proposed in [7], where systems of weakly interacting jump processes on time-varying random graphs with dynamically changing multi-color edges were studied. In [7], the dynamics of a node depend on the joint empirical distribution of all other nodes and the edges to which it connects. In contrast, the dynamics of an edge depend only on the corresponding nodes to which it connects. The paper [7] established the law of large numbers, propagation of chaos, and central limit theorems for these systems. Despite certain similarities, the class of models which we are considering in the current paper differs in several aspects from the models contemplated in [7]. First, the interacting particle systems that we study are on static block-structured graphs, whereas the ones considered in [7] are on time-varying random graphs with edge-structure dynamics. Moreover, in the current work, we consider a multi-population setting where the interaction between the nodes is local, i.e. each node interacts only with its neighbors, whereas in [7] the interaction between nodes is global, since the dynamic of a given node depends on the empirical distribution of all the other nodes. Finally, the analysis carried out and the results obtained in our current work are established on the vector of local empirical measures adapted to the multi-population context, and thus on product spaces, which allows us to overcome the heterogeneity due to the block structure of the graph. Furthermore, note that the current paper addresses the topic of interacting particle systems on large (random) networks, which has attracted increasing attention in recent years; see, e.g., [6, 7, 9] and the references therein.

Alongside the papers listed above, the multi-population framework has also been considered for systems of interacting diffusions. We mention for instance [47] for an analysis of a system of interacting Ornstein–Uhlenbeck processes on a heterogeneous network of credit-interlinked agents, [13, 16, 62] and the references therein for studies of neuronal networks composed of separate populations, or [58] and the references therein for mean-field multi-class interacting diffusions models in a general setting. (Note that some erroneous results that were originally stated in [62] were corrected in [63].)

The goal of the current work is to develop limiting results for interacting finite-state pure-jump processes on a class of block-structured networks. Our first main result, Theorem 5.1, and its consequence Corollary 5.1 give propagation of chaos and a law of large numbers under some regularity conditions on the degrees of the nodes. We show that in the mean-field limit, the asymptotic behavior of the node colors can be represented by the solution of a McKean–Vlasov system. Because of the lack of symmetry, we make use of the extension of the notion of chaoticity and Sznitman coupling methods to multi-class systems developed in [40, 43]. The existence and uniqueness results for the limiting system are established in Theorem 4.1. The regularity conditions which we impose (cf. Condition 4.1) can be compared to the uniform degree property introduced in [28] for a model of interacting diffusions on random graphs and to the one introduced in [15] for a model of interacting pure-jump processes on sparse graphs.

Another aspect which we are interested in is the large deviations properties of the system. For this purpose, with the aim of simplicity, we will restrict ourselves to the case where the blocks are cliques and the peripheral subgraph is complete, that is, the case where all peripheral nodes of the system are connected and all the central nodes within the same block are connected. We then state our next main results in Theorem 6.1, which establishes the large deviations principle for the *empirical measure vector* over finite time duration, followed by Theorem 6.2, which gives the large deviations principle for the *empirical process vector*. These results generalize those of [12, 49] to the multi-population context. Also, unlike [49] and similarly to [12], we do not impose chaotic initial conditions, but only converging initial conditions. The proofs of the large deviations principles, which provide tools for handling the technicalities arising from the multi-population context, generalize the classical approach developed in [24] and its adaptation to the context of jump processes in [49].

In summary, the current work is a contribution to the multi-population paradigm and a move towards heterogeneity for mean-field models and their large deviations behavior. The rest of this paper is organized as follows. The detailed model for interacting finite-state pure-jump processes on block-structured graphs is introduced in Section 2. Section 3 provides some practical examples of applications of the class of models studied in this paper. In Section 4, we introduce the McKean–Vlasov limiting system, and we prove the existence and uniqueness of its solution under specific regularity conditions introduced in Condition 4.1. Then, under the same conditions, in Section 5 we prove propagation of chaos (Theorem 5.1) and the law of large numbers (Corollary 5.1). Next, in Section 6, we present the large deviations principles for the empirical measure vector (Theorem 6.1) and for the empirical process vector (Theorem 6.2).

2. Formulation of the model

This section introduces the model and related notation.

2.1. The setting

A block-structured network:

- Consider an undirected block-structured graph $\mathcal{G} = (\mathcal{V}, \Xi)$, where \mathcal{V} is the set of nodes and Ξ is the set of edges. The set \mathcal{V} is partitioned into r (finite) blocks C_1, \dots, C_r of sizes N_1, \dots, N_r , respectively. Denote by $|\mathcal{V}| := N_1 + \dots + N_r = N$ the cardinality of the set \mathcal{V} , which corresponds to the total number of nodes in the network.
- The nodes of each block C_j are divided into two categories:

- **Central nodes** C_j^c are connected to some nodes from the same block but not to any nodes from any other blocks. We set $|C_j^c| = N_j^c$.
- **Peripheral nodes** C_j^p are connected to some nodes from the same block and some nodes from other blocks. We set $|C_j^p| = N_j^p$.

Multi-color nodes: Let $\mathcal{Z} := \{1, 2, \dots, K\} \subset \mathbb{N}$ be a set of K colors. Suppose that each node of the graph $\mathcal{G} = (\mathcal{V}, \Xi)$ is colored by one of these colors at each time. One can associate each node to a particle whose state space is \mathcal{Z} . Thus, we will use the denominations ‘node’ and ‘particle’ interchangeably to refer to the same thing. Denote by $(\mathcal{Z}, \mathcal{E})$ the directed graph where $\mathcal{E} \subset \mathcal{Z} \times \mathcal{Z} \setminus \{(z, z) | z \in \mathcal{Z}\}$ describes the set of admissible jumps for each particle. Moreover, whenever $(z, z') \in \mathcal{E}$, a particle colored by z is allowed to move from z to z' at a rate that depends on the current state of the node and the state of its neighbors (adjacent nodes).

For each $1 \leq j \leq r$ and $n \in C_j^c$ (resp. $n \in C_j^p$), let us define by $(X_{n,j}^c(t), t \geq 0)$ (resp. $(X_{n,j}^p(t), t \geq 0)$) the stochastic process that describes the state (color) of the central (resp. peripheral) node n at time t . In addition, we denote by $\mu_j^N(t)$ the local empirical measure describing the state of the j th block at time t , which is given by

$$\begin{aligned} \mu_j^N(t) &:= \frac{1}{N_j} \left(\sum_{n \in C_j^c} \delta_{X_{n,j}^c(t)} + \sum_{n \in C_j^p} \delta_{X_{n,j}^p(t)} \right) \\ &= \frac{N_j^c}{N_j} \frac{1}{N_j^c} \sum_{n \in C_j^c} \delta_{X_{n,j}^c(t)} + \frac{N_j^p}{N_j} \frac{1}{N_j^p} \sum_{n \in C_j^p} \delta_{X_{n,j}^p(t)} \\ &=: \frac{N_j^c}{N_j} \mu_j^{c,N}(t) + \frac{N_j^p}{N_j} \mu_j^{p,N}(t), \end{aligned} \tag{1}$$

where $\mu_j^{c,N}(t) = \frac{1}{N_j^c} \sum_{n \in C_j^c} \delta_{X_{n,j}^c(t)}$ (resp. $\mu_j^{p,N}(t) = \frac{1}{N_j^p} \sum_{n \in C_j^p} \delta_{X_{n,j}^p(t)}$) is the empirical measure describing the state of the central (resp. peripheral) nodes of the j th block at time t . The fractions $\frac{N_j^c}{N_j}$ (resp. $\frac{N_j^p}{N_j}$) thus represent the proportion of central (resp. peripheral) nodes in the block j . Denote by $\mathcal{M}_1(\mathcal{Z})$ the set of all probability measures over \mathcal{Z} , endowed with the topology of weak convergence.

The random dynamics: The process $X(t) = (X_{n,j}^c(t), X_{m,j}^p(t), n \in C_j^c, m \in C_j^p, 1 \leq j \leq r)$ describing the evolution of the entire system is a continuous-time Markov chain with state space \mathcal{Z}^N . The transition rate of each node depends on its current state and the state of its neighbors, together with the block to which it belongs. To characterize these neighborhoods, we introduce a set of local empirical measures describing the state of the star-shaped subgraph centered at each node n and composed of the nodes connected to it. To lighten the formulas and for ease of reading, we introduce the following shorthand notation: for any two nodes $n, m \in \mathcal{V}$, $m \sim n$ means that $\{m, n\} \in \Xi$. Moreover, for any block $1 \leq j \leq r$ and $\iota \in \{c, p\}$, we denote by $\mathfrak{N}_j^\iota(n) := \{n' \in C_j^\iota : n \sim n'\}$ the set of nodes in C_j^ι that are connected to n . Let $deg(n)$ denote the degree of the node n , and let $M_i^\iota(n) := |\mathfrak{N}_i^\iota(n)|$ for $\iota \in \{c, p\}$ and $1 \leq i \leq r$ be the cardinality of the set $\mathfrak{N}_i^\iota(n)$. Thus, one notices that for $n \in C_j^c$, $deg(n) = M_j^c(n) + M_j^p(n)$, and for $n \in C_j^p$, $deg(n) = M_j^c(n) + \sum_{k=1}^r M_k^p(n)$.

Now, for any $n \in C_j^c$ and $1 \leq j \leq r$, let us define

$$\aleph_{n,j}^c(t) := \frac{1}{1 + M_j^c(n)} \sum_{m \in \{n\} \cup \mathcal{N}_j^c(n)} \delta_{X_{m,j}^c(t)}, \quad \aleph_{n,j}^p(t) := \frac{1}{M_j^p(n)} \sum_{m \in \mathcal{N}_j^p(n)} \delta_{X_{m,j}^p(t)}, \tag{2}$$

and

$$\varrho_{n,j}^c := \frac{1 + M_j^c(n)}{1 + \text{deg}(n)}, \quad \varrho_{n,j}^p := \frac{M_j^p(n)}{1 + \text{deg}(n)}, \tag{3}$$

and finally

$$\mu_{n,j}^{c,N}(t) := \varrho_{n,j}^c \aleph_{n,j}^c(t) + \varrho_{n,j}^p \aleph_{n,j}^p(t). \tag{4}$$

Equivalently, for any $n \in C_j^p$, $1 \leq j \leq r$, and $j' \neq j$, define

$$\begin{aligned} \beth_{n,j}^c(t) &:= \frac{1}{M_j^c(n)} \sum_{m \in \mathcal{N}_j^c(n)} \delta_{X_{m,j}^c(t)}, & \beth_{n,j,j'}^p(t) &:= \frac{1}{1 + M_j^p(n)} \sum_{m \in \{n\} \cup \mathcal{N}_j^p(n)} \delta_{X_{m,j}^p(t)}, \\ \beth_{n,j,j'}^p(t) &:= \frac{1}{M_j^p(n)} \sum_{m \in \mathcal{N}_j^p(n)} \delta_{X_{m,j'}^p(t)}, \end{aligned} \tag{5}$$

and

$$\varsigma_{n,j}^c := \frac{M_j^c(n)}{1 + \text{deg}(n)}, \quad \varsigma_{n,j,j'}^p := \frac{1 + M_j^p(n)}{1 + \text{deg}(n)}, \quad \varsigma_{n,j,j'}^p := \frac{M_j^p(n)}{1 + \text{deg}(n)}, \tag{6}$$

and finally

$$\mu_{n,j}^{p,N}(t) := \varsigma_{n,j}^c \beth_{n,j}^c(t) + \sum_{\substack{j'=1 \\ j' \neq j}}^r \varsigma_{n,j,j'}^p \beth_{n,j,j'}^p(t). \tag{7}$$

Therefore, the random dynamic in each block $1 \leq j \leq r$ is described as follows:

- **The central nodes dynamic.** For each central node $n \in C_j^c$, its color $X_{n,j}^c(t)$ goes from z to z' , for $(z, z') \in (\mathcal{Z}, \mathcal{E})$, at rate

$$\lambda_{j,z,z'}^c \left(\aleph_{n,j}^c(t), \aleph_{n,j}^p(t), \varrho_{n,j}^c, \varrho_{n,j}^p \right), \tag{8}$$

which depends on its current state and on the states of its neighbors through the functions $\lambda_{j,z,z'}^c : \mathcal{M}_1(\mathcal{Z}) \times \mathcal{M}_1(\mathcal{Z}) \times [0, 1] \times [0, 1] \rightarrow \mathbb{R}_+$.

- **The peripheral nodes dynamic.** For each peripheral node $n \in C_j^p$, its color $X_{n,j}^p(t)$ goes from z to z' , for $(z, z') \in (\mathcal{Z}, \mathcal{E})$, at rate

$$\lambda_{j,z,z'}^p \left(\beth_{n,j}^c(t), \beth_{n,j,1}^p(t), \dots, \beth_{n,j,r}^p(t), \varsigma_{n,j}^c, \varsigma_{n,j,1}^p, \dots, \varsigma_{n,j,r}^p \right), \tag{9}$$

which also depends on its state and the states of its neighbors through the functions $\lambda_{j,z,z'}^p : (\mathcal{M}_1(\mathcal{Z}))^{r+1} \times [0, 1]^{r+1} \rightarrow \mathbb{R}_+$.

The explicit forms of the rate functions will be introduced in Condition 4.1. To avoid cluttering our notation, let us introduce the following vectors:

$$v_{n,j}^{c,N}(t) := \left(\mathfrak{K}_{n,j}^c(t), \mathfrak{K}_{n,j}^p(t), \varrho_{n,j}^c, \varrho_{n,j}^p \right), \tag{10}$$

$$v_{n,j}^{p,N}(t) := \left(\mathfrak{Z}_{n,j}^c(t), \mathfrak{Z}_{n,j,1}^p(t), \dots, \mathfrak{Z}_{n,j,r}^p(t), \mathfrak{S}_{n,j}^c, \mathfrak{S}_{n,j,1}^p, \dots, \mathfrak{S}_{n,j,r}^p \right). \tag{11}$$

Thus, we will write $\lambda_{j,z,z'}^c(v_{n,j}^{c,N}(t))$ instead of (8) and $\lambda_{j,z,z'}^p(v_{n,j}^{p,N}(t))$ instead of (9).

Remark 2.1. One can see the model under investigation as a multi-species system where each block C_j represents a separate species. In particular, the rate functions $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ being block-dependent, the dynamic of each particle depends on its species, i.e., the block to which it belongs. This idea has been extensively used in the literature on multi-type systems; see, e.g., [1, 3, 5, 16] and the references therein. The specificity here is the existence of heterogeneity even across particles of the same species/block. Indeed, the central/peripheral paradigm creates two sub-types of particles within the same species whose rate functions differ. This construction appears to be natural in certain multi-group systems where only a few particles from the different groups interact; detailed examples are given below. Also, the interaction structure differs even for the central (resp. peripheral) particles of the same species/block, given that the rate functions depend on the node-centered local empirical measures, which differ even within the same block.

2.2. The infinitesimal generator

For any $T \in (0, +\infty)$, the processes $X_{n,j}^c : [0, T] \rightarrow \mathcal{Z}$ for $n \in C_j^c$ and $X_{m,j}^p : [0, T] \rightarrow \mathcal{Z}$ for $m \in C_j^p$, which respectively describe the evolution of the central and the peripheral particles over the time interval $[0, T]$, are càdlàg paths, and thus are elements of the Skorokhod space $\mathcal{D}([0, T], \mathcal{Z})$ equipped with the Skorokhod topology. Let $X^N = (X_{n,j}^c, X_{m,j}^p, n \in C_j^c, m \in C_j^p, 1 \leq j \leq r) \in \mathcal{D}([0, T], \mathcal{Z}^N)$ denote the full path description of all N particles. Thus the process X^N is a Markov process with càdlàg paths, with state space \mathcal{Z}^N , and with the infinitesimal generator \mathcal{L}^N acting on the bounded measurable functions ϕ on \mathcal{Z}^N according to

$$\begin{aligned} \mathcal{L}^N \phi(x^N) := & \sum_{j=1}^r \left[\sum_{n \in C_j^c} \sum_{z', z \in \mathcal{E}} \lambda_{j,z,z'}^c \right. \\ & \left. \left(\frac{1}{M_j^c(n) + 1} \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \delta_{x_{m,j}}, \frac{1}{M_j^p(n)} \sum_{m \in \mathfrak{N}_j^p(n)} \delta_{x_{m,j}}, \varrho_{n,j}^c, \varrho_{n,j}^p \right) \right. \\ & \left. \times \left(\phi(x_{n,z,z'}^N) - \phi(x^N) \right) \right. \\ & + \sum_{n \in C_j^p} \sum_{z', z \in \mathcal{E}} \lambda_{j,z,z'}^p \left(\frac{1}{M_j^c(n)} \sum_{m \in \mathfrak{N}_j^c(n)} \delta_{x_{m,j}}, \frac{1}{M_1^p(n)} \sum_{m \in \mathfrak{N}_1^p(n)} \delta_{x_{m,1}}, \dots, \right. \\ & \left. \frac{1}{M_j^p(n) + 1} \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \delta_{x_{m,j}}, \dots \right) \end{aligned}$$

$$\left. \begin{aligned} &\dots, \frac{1}{M_r^p(n)} \sum_{m \in \mathcal{N}_r^p(n)} \delta_{x_{m,r}, \varsigma_{n,j}^c, \varsigma_{n,j,1}^p, \dots, \varsigma_{n,j,r}^p} \\ &\quad \times \left(\phi(x_{n,z,z'}^N) - \phi(x^N) \right) \end{aligned} \right],$$

where $x^N = (x_{n,j}, x_{m,j}, n \in C_j^c, m \in C_j^p, 1 \leq j \leq r) \in \mathcal{Z}^N$ and $x_{n,z,z'}^N$ describes the new configuration of the system when the state of the n th node has changed from z to z' .

2.3. Stochastic differential equation representation

Recall that, for each central node $n \in C_j^c$ (resp. peripheral node $n \in C_j^p$) from a given block $1 \leq j \leq r$, the evolution of its color is described by the continuous-time stochastic process $(X_{n,j}^c(t), t \geq 0)$ (resp. $(X_{n,j}^p(t), t \geq 0)$) that takes values in the finite state space \mathcal{Z} , and whose dynamic is given by the time-dependent transition rate matrix $\left(\lambda_{j,z,z'}^c(v_{n,j}^{c,N}(t)) \right)_{(z,z') \in \mathcal{E}}$ (resp. $\left(\lambda_{j,z,z'}^p(v_{n,j}^{p,N}(t)) \right)_{(z,z') \in \mathcal{E}}$). Therefore, using a classical approach (see e.g. [60, p. 104]), the processes $X_{n,j}^c$ and $X_{n,j}^p$ can be represented, at least weakly, by the following system of stochastic differential equations:

$$\begin{aligned} X_{n,j}^c(t) &= X_{n,j}^c(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{X_{n,j}^c(s-)=z(z'-z)} \mathbb{1}_{\left[0, \lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s-))\right]}(y) \mathcal{N}_{n,j}^c(ds, dy), \\ X_{n,j}^p(t) &= X_{n,j}^p(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{X_{n,j}^p(s-)=z(z'-z)} \mathbb{1}_{\left[0, \lambda_{j,z,z'}^p(v_{n,j}^{p,N}(s-))\right]}(y) \mathcal{N}_{n,j}^p(ds, dy), \end{aligned} \tag{12}$$

where $\{\mathcal{N}_{n,j}^c, n \in C_j^c, 1 \leq j \leq r\}$ and $\{\mathcal{N}_{n,j}^p, n \in C_j^p, 1 \leq j \leq r\}$ are collections of Poisson random measures on \mathbb{R}^2 whose intensity measures are Lebesgue on \mathbb{R}_+^2 . We will use the representation (12) in the analysis of the asymptotic behavior of the system when the total number of nodes N goes to infinity.

3. Examples

As mentioned in the introduction, mean-field block models have been proposed to investigate various phenomena arising in fields such as physics, engineering, biology, etc. This section presents some examples of applications of the model analyzed in the current paper, with the goal being to illustrate its usefulness and its flexibility to capture various phenomena. Of course, it remains a toy model that should be appropriately adapted to different applications, but we believe that the insights from the current study are of great interest for both theoretical and practical purposes.

3.1. Load-balancing networks

Load-balancing protocols are often used in queueing networks to improve system performance by shortening the queue length, reducing the waiting time, and increasing the system throughput. In this regard, the mean-field approach has been proven to be useful; see, e.g., [55, 65, 66]. In particular, interesting work in this direction was proposed in [26], where the authors considered a queueing network with N nodes in which queue lengths are balanced

through mean-field interactions using an *interaction function*. Here we summarize their model and then describe how our current model can be used to generalize the ideas in [26].

Consider a system consisting of N queues with a mean-field interaction. At $t = 0$, for $1 \leq n \leq N$, the arrival rate to the n th queue occurs according to $\zeta_{X_n(0)}$, and the service rate at queue n is $\vartheta_{X_n(0)}$. Let $h(x) : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ be a continuous nondecreasing *interaction function* satisfying certain regularity conditions (see [26, p. 339]). This function makes it possible to capture the mean-field interaction between queues as follows: for each queue $n = 1, 2, \dots, N$, the arrival rate at time t is given by $\zeta_{X_j(t)} - h(X_j(t), \langle \mu^N(t)(dx), x \rangle)$, where $\mu^N(t) := \frac{1}{N} \sum_{j=1}^N \delta_{X_j(t)}$ is the empirical measure corresponding to the N queues at time t , and $\langle \mu^N(t)(dx), x \rangle = \frac{1}{N} \sum_{j=1}^N X_j(t)$ is the mean queue length of the N queues at time t . Roughly speaking, the arrival rate at each queue depends on the current size of the queue and on the mean size of its neighbors (which is the entire set of queues in this case). The authors of [26] studied the performance of this system when the number of queues N goes to infinity.

The model proposed in the current paper can be seen as a generalization of the model in [26] to heterogeneous queueing networks, namely, to block-structured networks. To see this, let us consider the graph $\mathcal{G} = (\mathcal{V}, \Xi)$ as a queueing network where the particles (nodes) are finite-buffer server queues of maximum size K (arbitrarily large), and the corresponding states $(X_{n,j}^c(t), X_{m,j}^p(t), n \in C_j^c, m \in C_j^p, 1 \leq j \leq r, t \geq 0)$ represent the number of customers waiting in each queue at each time t . Again, at $t = 0$, for $1 \leq n \leq N$, the arrival rate to the n th queue occurs according to $\zeta_{X_{n,j}^c(0)}$, and the service rate at queue n is $\vartheta_{X_{n,j}^c(0)}$, for $\iota \in \{c, p\}$. Since the network now is heterogeneous, the mean-field interaction is local. Thus, the arrival rate at a central node queue $n \in C_j^c$ at time t is given by $\zeta_{X_{n,j}^c(t)} - h^c(X_{n,j}^c(t), \langle \mu_{n,j}^{c,N}(t)(dx), x \rangle)$, whereas the arrival rate at a peripheral node queue $n \in C_j^p$ at time t is given by $\zeta_{X_{n,j}^p(t)} - h^p(X_{n,j}^p(t), \langle \mu_{n,j}^{p,N}(t)(dx), x \rangle)$, where $\mu_{n,j}^{c,N}(t)$ and $\mu_{n,j}^{p,N}(t)$ are the local empirical measures respectively given by (4) and (7). The service rates $\vartheta_{X_{n,j}^c(t)}$ and $\vartheta_{X_{n,j}^p(t)}$ depend only on the queue sizes $X_{n,j}^c(t)$ and $X_{n,j}^p(t)$ at time t . Hence, the transition rates $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ are specified as follows:

- The size $X_{n,j}^c(t)$ of each central queue $n \in C_j^c$ at time t goes from z to z' at rate

$$\lambda_{j,z,z'}^c := \begin{cases} \zeta_{X_{n,j}^c(t)} - h^c\left(X_{n,j}^c(t), \frac{1}{\text{deg}(n)} \sum_{\iota \in \{c,p\}} \sum_{m \in \mathfrak{N}_k^\iota(n)} X_{m,j}^m(t)\right) & \text{if } z' = z + 1 \text{ and } z' \leq K, \\ \vartheta_{X_{n,j}^c(t)} & \text{if } z' = z - 1 \text{ and } X_{n,j}^c(t) \geq 1, \\ - \sum_{y \neq z} \lambda_{j,z,y}^c & \text{if } z' = z, \\ 0 & \text{otherwise.} \end{cases}$$

- The size $X_{n,j}^p(t)$ of each peripheral queue $n \in C_j^p$ at time t goes from z to z' at rate

$$\lambda_{j,z,z'}^p := \begin{cases} \zeta_{X_{n,j}^p(t)} - h^p\left(X_{n,j}^p(t), \frac{1}{\text{deg}(n)} \sum_{\substack{1 \leq k \leq r \\ \iota \in \{c,p\}}} \sum_{m \in \mathfrak{N}_k^\iota(n)} (X_{m,k}^\iota(t))\right) & \text{if } z' = z + 1 \text{ and } z' \leq K, \\ \vartheta_{X_{n,j}^p(t)} & \text{if } z' = z - 1 \text{ and } X_{n,j}^p(t) \geq 1, \\ - \sum_{y \neq z} \lambda_{j,z,y}^p & \text{if } z' = z, \\ 0 & \text{otherwise.} \end{cases}$$

It is worth mentioning that the sparse graph topologies have been considered in applications in response to some issues encountered when implementing load-balancing protocols. In particular, many service systems are geographically constrained; therefore, when a task arrives at any specific server, it may be impossible to collect instantaneous state information from all the servers. In addition, executing a task commonly involves the use of some data, and storing such data for all possible tasks on all servers requires an excessive amount of storage capacity. The use of sparser graph topologies is then considered, such that tasks that arrive at a specific server can only be forwarded, following a specific load-balancing scheme, to the servers that possess the data required to process the tasks. In other words, a specific server can only interact with its neighbors in a suitable sparse topology; see, e.g., [15] and the references therein for more insights about the subject. The block-structured topology with the central/peripheral paradigm can for instance be considered to overcome the geographic constraint by allowing central nodes to rely only on the information collected locally on nodes from the same block, which may represent nodes within the same geographic area, while the peripheral nodes are those relying on information from both within and outside the block. To increase system efficiency, one could restrict the number of peripheral nodes allowed. The results obtained in the current work allow us to understand the behavior of such systems when the number N of servers of the network is very large. In particular, the multi-chaotic property established in Theorem 5.1 tells us that the queue lengths at any finite collection of tagged servers are statistically asymptotically independent, and the queue-length process for each server converges in distribution to the corresponding McKean–Vlasov process given by (14). Also, Condition 4.1 and Remark 4.1 tell us that the multi-chaoticity result holds even when the peripheral subgraph is not complete, which means that one can achieve similar asymptotic performance even with far fewer connections between the peripheral nodes than when all the peripheral nodes are connected and all the central nodes of the same block are connected.

3.2. Multi-population SIS epidemics

The susceptible–infected–susceptible (SIS) model, originally used in epidemiology, is also convenient to model the spread of information in networks, since the two phenomena are similar. The SIS model can be summarized as follows: consider a piece of information, or an infectious disease, that propagates across a population. A member that has a copy of the information/disease is said to be *infected*, and a member that does not have a copy of the information/disease is said to be *susceptible*. When an infected member comes into contact with a susceptible one, the former transmits a copy of the information (disease) to the latter, which gets infected. Moreover, an infected member may spontaneously get rid of the information/disease, a phenomenon called *curing*, and become susceptible again.

In both epidemiology and network information diffusion, the population often consists of relatively isolated subgroups such that members of the same subgroup interact a great deal, but only a few pairs of members from different subgroups are connected. One might think, for instance, of countries as isolated communities connected by travelers across the globe, or of interactions in social media, which often happen in almost closed communities, with only a few influential members interacting across groups. Our model allows one to study the spreading dynamics of information or of a disease among the members of a population structured as separate communities.

Consider a population consisting of r isolated communities and a ‘mobile’ community. The members of each isolated community interact only among themselves and with members of the mobile community. Thus, there is no direct interaction between members of different

communities. However, indirect inter-community interactions happen via the set of mobile members. This idea was used in [2], where the authors considered an optimal control problem to find the optimal resource allocation strategy that maximizes information spread over a multi-community population. Their objective was to obtain a good tradeoff between the information spread in the network and the use of system resources.

Now, let $\mathcal{Z} := \{0, 1\}$ be the state space that indicates whether the particle is *susceptible* ($= 0$) or *infected* ($= 1$). Recalling the model description introduced in Section 2, one might think of the central nodes of each block as an isolated community that interacts with other communities only through the members of the mobile community represented by the peripheral nodes. Note that in contrast to [2], the central nodes of a given block interact only with peripheral (mobile) nodes from the same block, and not with all the peripheral/mobile nodes, as stipulated in [2]. Also, the interaction graph for the peripheral members is not complete; thus, not all the peripheral nodes interact with each other. Nevertheless, the fact that the multi-chaotic property holds under Condition 4.1 (cf. Theorem 5.1) tells us that systems with full connections among the peripheral components and among the components of each block are asymptotically close to those with fewer connections, as specified by Condition 4.1 and Remark 4.1. This is of interest, for example, in resource allocation problems where a cost is attributed to each connection; however, such considerations are beyond the scope of the present paper.

Denote by $X_{n,j}^c(t)$, for $n \in C_j^c$, and $X_{m,j}^p(t)$, for $m \in C_j^p$, the state ('susceptible' or 'infected') of the n th central particle and the m th peripheral particle, respectively, in the j th community. Two connected central members of the same community j come in contact with each other at rate γ_j . Connected peripheral and central nodes from the same community interact with each other at rate ν_j . Two connected peripheral nodes come in contact with each other at rate η . Finally, an infected node in the j th community spontaneously gets rid of the infection at rate ζ_j . Therefore, the transition rates $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$, which sum up the dynamics that we are interested in, are specified as follows:

- The state $X_{n,j}^c(t)$ of each central member $n \in C_j^c$ at time t goes from z to z' at rate

$$\lambda_{j,z,z'}^c := \begin{cases} \sum_{m \in \mathfrak{N}_j^c(n)} X_{m,j}^c(t)\gamma_j + \sum_{m \in \mathfrak{N}_j^p(n)} X_{m,j}^p(t)\nu_j & \text{if } z = 0 \text{ and } z' = 1, \\ \zeta_j & \text{if } z = 1 \text{ and } z' = 0, \\ - \sum_{y \neq z} \lambda_{j,z,y}^c & \text{if } z' = z, \\ 0 & \text{otherwise.} \end{cases}$$

- The state $X_{n,j}^p(t)$ of each peripheral (mobile) member $n \in C_j^p$ at time t goes from z to z' at rate

$$\lambda_{j,z,z'}^p := \begin{cases} \sum_{m \in \mathfrak{N}_j^c(n)} X_{m,j}^c(t)\nu_j + \sum_{k=1}^r \sum_{m \in \mathfrak{N}_k^p(n)} X_{m,k}^p(t)\eta & \text{if } z = 0 \text{ and } z' = 1, \\ \zeta_j & \text{if } z = 1 \text{ and } z' = 0, \\ - \sum_{y \neq z} \lambda_{j,z,y}^p & \text{if } z' = z, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the large deviations properties established in Section 6 constitute a step towards the study of the large-time behavior of such systems. Indeed, the large deviations of the empirical measures established in Theorem 6.2 can be used to investigate the large deviations of the

invariant measure, from which one can study the large-time behavior of the system and related phenomena such as metastability and convergence to the invariant measure. The interested reader can consult, e.g., [25, 36, 44, 67] and the references therein for further insight.

4. Existence and uniqueness of the limiting system

This section aims to introduce and prove the existence of the limiting equation that describes the behavior of the interacting particle system detailed in Section 2, as the total number of particles N in the system tends to infinity. In particular, this equation is of McKean–Vlasov type as explained below. The main result of this section is Theorem 4.1, which establishes the existence and uniqueness of the solution of the limiting McKean–Vlasov equation. The convergence of the system towards this equation will then be investigated in Section 5.

4.1. Notation and conventions

Let (\mathbb{S}, d) be a Polish space. For any $y \in \mathbb{S}^d$, for some $d \in \mathbb{N}$, one writes $\|y\| := \max(y_1, \dots, y_d)$. For any $x \in \mathcal{D}([0, T], \mathbb{S}^d)$, $\|x\|_T$ denotes $\sup_{0 \leq t \leq T} \|x(t)\|$. Let $\mathcal{M}(\mathbb{S})$ be the set of all measures on \mathbb{S} . Given $\mu, \nu \in \mathcal{M}(\mathbb{S})$, the bounded-Lipschitz metric $d_{BL}(\cdot, \cdot)$ is defined by

$$d_{BL}(\mu, \nu) := \sup_{g \in Lip(\mathbb{S})} |\langle \mu, g \rangle - \langle \nu, g \rangle|, \tag{13}$$

where

$$Lip(\mathbb{S}) := \left\{ g \in C_b(\mathbb{S}) : \sup_{x \in \mathbb{S}} |g(x)| \leq 1, \sup_{x \neq y} \frac{|g(x) - g(y)|}{d(x, y)} \leq 1 \right\}.$$

Recall that the bounded-Lipschitz metric metrizes the weak convergence of probability measures on \mathbb{S} with respect to bounded continuous test functions $C_b(\mathbb{S})$. For $p \geq 1$, let $\mathcal{P}_p(\mathbb{S})$ be the collection of all probability measures on \mathbb{S} with finite p th moment. Then, for any μ and ν in $\mathcal{P}_p(\mathbb{S})$, the p th Wasserstein distance between μ and ν is defined as

$$\mathcal{W}_p(\mu, \nu) := \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{E \times E} d(x, y)^p d\gamma(x, y) \right)^{1/p},$$

where $\Gamma(\mu, \nu)$ denotes the collection of all measures on $\mathbb{S} \times \mathbb{S}$ with marginals μ and ν . Moreover, for M_1, M_2 in $\mathcal{P}_p(\mathcal{D}([0, T], \mathbb{S}) \times \dots \times \mathcal{D}([0, T], \mathbb{S}))$, the p th Wasserstein distance between M_1 and M_2 is given, for any $t \in [0, T]$, by

$$\mathcal{W}_{p,t}(M_1, M_2) := \inf \left\{ [\mathbb{E} \|Y_1 - Y_2\|_t^p]^{1/p} : Y_1, Y_2 \in \mathcal{D}([0, T], \mathbb{S}) \times \dots \times \mathcal{D}([0, T], \mathbb{S}), M_1 = \mathcal{L}(Y_1), M_2 = \mathcal{L}(Y_2) \right\}.$$

4.2. The limiting system

We use in the sequel the convention that N goes to infinity when both $\min_{1 \leq j \leq r} N_j^c$ and $\min_{1 \leq j \leq r} N_j^p$ go to infinity. Given the multi-population setting, one describes the state of the system at each time t using the following empirical measure vector:

$$\mu^N(t) := \left(\mu_1^{c,N}(t), \mu_1^{p,N}(t), \dots, \mu_r^{c,N}(t), \mu_r^{p,N}(t) \right),$$

where we recall that for each $1 \leq j \leq r$, $\mu_j^{c,N}(t)$ (resp. $\mu_j^{p,N}(t)$) is the empirical measure describing the states of the central (resp. peripheral) nodes of the j th block at time t . Under some regularity conditions (cf. Condition 4.1), we will prove in Section 5 the convergence, as N tends to infinity, of the empirical measure vector μ^N towards the distribution $\mu \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$, the solution of an appropriate limiting system. Namely, the empirical vector μ^N should converge weakly to μ with

$$\mu := (\mu_1^c, \mu_1^p, \dots, \mu_r^c, \mu_r^p) \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r},$$

where $\mu_j^c := \mathcal{L}(\bar{X}_{n,j}^c)$ for $n \in C_j^c$ and $\mu_j^p := \mathcal{L}(\bar{X}_{m,j}^p)$ for $m \in C_j^p$, with $(\bar{X}_{n,j}^c, \bar{X}_{m,j}^p, n \in C_j^c, m \in C_j^p, 1 \leq j \leq r)$ being the solution of the following system of stochastic differential equations:

$$\begin{aligned} \bar{X}_{n,j}^c(t) &= \bar{X}_{n,j}^c(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{\bar{X}_{n,j}^c(s-) = z} (z' - z) \mathbb{1}_{[0, \lambda_{j,z,z'}^c]} (v_j^c(s-))^{(y)} \mathcal{N}_{n,j}^c(ds, dy), \\ \bar{X}_{m,j}^p(t) &= \bar{X}_{m,j}^p(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{\bar{X}_{m,j}^p(s-) = z} (z' - z) \mathbb{1}_{[0, \lambda_{j,z,z'}^p]} (v_j^p(s-))^{(y)} \mathcal{N}_{m,j}^p(ds, dy). \end{aligned} \tag{14}$$

The vectors $v_j^c(t)$ and $v_j^p(t)$ are defined by

$$\begin{aligned} v_j^c(t) &:= (\mu_j^c(t), \mu_j^p(t), p_j^c, p_j^p), \\ v_j^p(t) &:= (\mu_j^c(t), \mu_1^p(t), \dots, \mu_r^p(t), \alpha_j^c, q_{j,1}, \dots, q_{j,r}), \end{aligned} \tag{15}$$

where $p_j^c, p_j^p, \alpha_j^c, q_{j,1}, \dots, q_{j,r} \in (0, 1)$ are parameters satisfying

$$p_j^c + p_j^p = 1 \quad \text{and} \quad \alpha_j^c + q_{j,1} + \dots + q_{j,r} = 1 \quad \text{for each } 1 \leq j \leq r;$$

these parameters will later be chosen appropriately (cf. Condition 4.1). The link between the initial conditions of the systems (12) and (14) will be introduced in the sequel. Observe that the solution of (14) depends not only on its sample path but also on the distribution of the process itself. Thus, the system (14) is McKean–Vlasov.

4.3. Regularity assumptions

We introduce and discuss here the regularity conditions under which the existence and uniqueness of the limiting system (14), together with the propagation of chaos and laws of large numbers investigated in Section 5, hold.

Condition 4.1.

1. For all $1 \leq j \leq r$ and $(z, z') \in \mathcal{E}$, there exist measurable functions $\gamma_{z,z'}^{j,c} : \mathcal{Z} \rightarrow \mathbb{R}^+$ and $\gamma_{z,z'}^{j,p} : \mathcal{Z} \rightarrow \mathbb{R}^+$ such that the following hold:

- For any probability measures $\nu, \mu \in \mathcal{M}_1(\mathcal{Z})$ and any real numbers a_1, a_2 satisfying $0 < a_1, a_2 < 1$ and $a_1 + a_2 = 1$, we have

$$\lambda_{j,z,z'}^c(\nu, \mu, a_1, a_2) = a_1 \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \nu(dx) + a_2 \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu(dx). \tag{16}$$

- For any $\nu, \mu_1, \dots, \mu_r \in \mathcal{M}_1(\mathcal{Z})$ and any real numbers a, b_1, \dots, b_r satisfying $0 < a, b_1, \dots, b_r < 1$ and $a + b_1 + \dots + b_r = 1$, we have

$$\lambda_{j,z,z'}^p(\nu, \mu_1, \dots, \mu_r, a, b_1, \dots, b_r) = a \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \nu(dx) + b_1 \int_{\mathcal{Z}} \gamma_{z,z'}^{1,p}(x) \mu_1(dx) + \dots + b_r \int_{\mathcal{Z}} \gamma_{z,z'}^{r,p}(x) \mu_r(dx). \tag{17}$$

2. For each block $1 \leq j \leq r$, there exist $p_j^c, p_j^p \in (0, 1)$ such that, as $N \rightarrow \infty$,

$$\frac{N_j^p}{N_j} \rightarrow p_j^p, \quad \frac{N_j^c}{N_j} \rightarrow p_j^c, \quad \text{and} \quad p_j^p + p_j^c = 1. \tag{18}$$

3. For each block $1 \leq j \leq r$, as $N \rightarrow \infty$,

$$\sup_{n \in C_j^c} |q_{n,j}^c - p_j^c| \rightarrow 0 \quad \text{and} \quad \sup_{n \in C_j^p} |q_{n,j}^p - p_j^p| \rightarrow 0. \tag{19}$$

4. For each block $1 \leq j \leq r$, there exist $\alpha_j^c, q_{j1}, \dots, q_{jr} \in (0, 1)$ with $\alpha_j^c + q_{j1} + \dots + q_{jr} = 1$ such that the following conditions hold for each block $1 \leq i \leq r$, as $N \rightarrow \infty$:

$$\sup_{n \in C_j^p} |s_{n,j}^c - \alpha_j^c| \rightarrow 0 \quad \text{and} \quad \sup_{n \in C_j^p} |s_{n,j,i}^p - q_{ji}| \rightarrow 0. \tag{20}$$

5. For all nodes $n \in \mathcal{V}$, $\text{deg}(n) \rightarrow \infty$ as $N \rightarrow \infty$.

Remark 4.1.

1. Since \mathcal{Z} is a finite state space, the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ are bounded on \mathcal{Z} . Moreover, since $\mathcal{Z} \subset \mathbb{N}$ and since every bounded function on \mathbb{N} is automatically Lipschitz, $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ are also Lipschitz. Denote by $\bar{\gamma} > 0$ the maximum bound and by L_γ the maximum Lipschitz coefficient of the sequences of functions $\{\gamma_{z,z'}^{j,c}, 1 \leq j \leq r, (z, z') \in \mathcal{E}\}$ and $\{\gamma_{z,z'}^{j,p}, 1 \leq j \leq r, (z, z') \in \mathcal{E}\}$.
2. The conditions in (19) and (20) are satisfied if, for instance, for all $1 \leq j \leq r$, the following hold:
 - For each $n \in C_j^c$, we have $M_j^c(n)/N_j^c \rightarrow 1$ and $M_j^p(n)/N_j^p \rightarrow 1$ as $N \rightarrow \infty$.
 - For each $n \in C_j^p$, we have $M_j^c(n)/N_j^c \rightarrow 1$ and $M_i^p(n)/N_i^p \rightarrow 1$ as $N \rightarrow \infty$ for all $1 \leq i \leq r$.

Indeed, under this assumption one can define

$$\alpha_j^c = \lim_{N \rightarrow \infty} \frac{N_j^c}{N_j^c + N_1^p + \dots + N_r^p} \quad \forall 1 \leq j \leq r, \tag{21}$$

$$q_{ji} = \lim_{N \rightarrow \infty} \frac{N_i^p}{N_j^c + N_1^p + \dots + N_r^p} \quad \forall 1 \leq j, i \leq r, \tag{22}$$

and thus, one can easily verify that, as $N \rightarrow \infty$, the following hold:

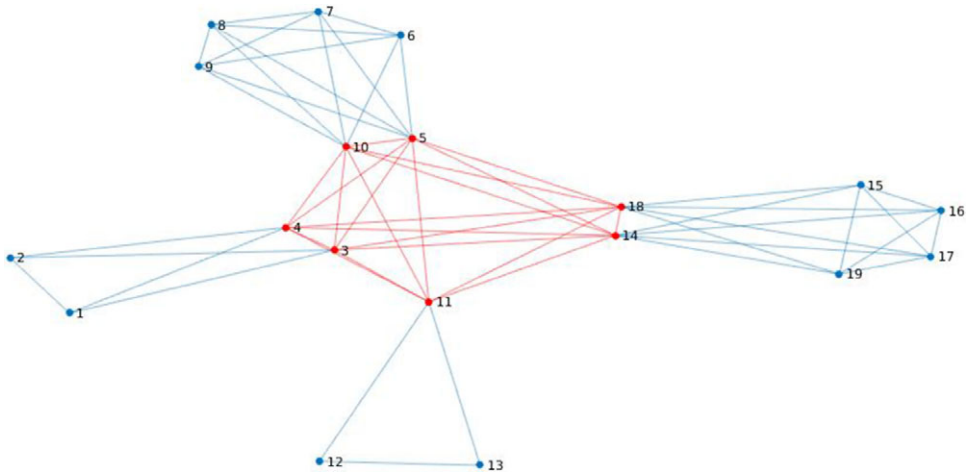


FIGURE 1. Example of a block-structured graph with a complete peripheral subgraph. Here we have a 4-block-structured graph linked by a set of peripheral nodes. For the first block, the set of central nodes is $C_1^c = \{1, 2\}$ and the set of peripheral nodes is $C_1^p = \{3, 4\}$. The set of all peripheral nodes of the graph is given by the set of nodes $C^p = \{3, 4, 5, 10, 11, 14, 18\}$.

- For all $n \in C_j^c$,

$$\frac{1 + M_j^c(n)}{1 + \text{deg}(n)} \rightarrow p_j^c \text{ and } \frac{M_j^p(n)}{1 + \text{deg}(n)} \rightarrow p_j^p. \tag{23}$$

- For all $n \in C_j^p$ and $i \neq j$,

$$\frac{M_i^p(n)}{1 + \text{deg}(n)} \rightarrow q_{ji}, \quad \frac{1 + M_j^p(n)}{1 + \text{deg}(n)} \rightarrow q_{jj}, \quad \text{and} \quad \frac{M_j^c(n)}{1 + \text{deg}(n)} \rightarrow \alpha_j^c. \tag{24}$$

3. A special case where the conditions (19) and (20) are satisfied is when the blocks are cliques and the peripheral subgraph is complete—that is, when all peripheral nodes are connected (see Figure 1) and all the nodes in the same block are connected. In such a case, the central (resp. peripheral) nodes in the same block are exchangeable.
4. Even though the conditions (19) and (20) are somewhat restrictive, the construction of the model allows one to have very different degrees in each block. One might further compare these conditions with existing conditions in the literature. Consider for example the condition imposed in [15] for a supermarket model on sparse graphs to asymptotically behave as on cliques. The condition in [15] relies on the local properties of the graph by requiring direct neighbors of any node to have asymptotically similar degrees; see [15, Condition 1(ii)]. This condition is violated in our model. Indeed, the conditions (19) and (20) allow central and peripheral nodes from the same block to have very different degrees, even if they are neighbors, which goes beyond [15, Condition 1(ii)]. In addition, under our condition, $\text{deg}_{\max}(G)/\text{deg}_{\min}(G)$ should not go to 1 as $N \rightarrow \infty$, nor does $\max_j |(\text{deg}_{\min}(C_j)/\text{deg}_{\max}(C_j)) - 1|$ go to zero as proposed in [15, Remark 1] (here $\text{deg}_{\min}(C_j)$ and $\text{deg}_{\max}(C_j)$ refer to the minimum and maximum degrees of nodes within the same block j). In this sense, the family of graphs which we are considering

in the present work is sparser than the ones covered by [15, Condition 1(ii)]. Another condition with which to compare ours is the one proposed in [28], under which an n -dimensional diffusion system converges to a limiting Fokker–Planck equation; see [28, Equations (1.1) and (1.3)]. Note that [28, Equations (1.5) and (1.7)] impose global regularity conditions in the sense that the degrees of all the nodes should converge to the same limit; such conditions are not imposed here.

5. While the current paper considers deterministic graphs, one can investigate the case where the underlying graph topology is random. For example, it is of interest for some applications to have a scenario where the connections between the peripheral nodes are random. One can then search for adequate conditions to impose on the edge dynamics for the propagation-of-chaos property to hold. This, however, goes beyond the scope of the current paper.

4.4. Existence and uniqueness

We now prove the existence and uniqueness of the solution of the limiting McKean–Vlasov system introduced in (14).

Theorem 4.1. *Suppose that Condition 4.1 holds. Then, for a given initial condition $\left(\left(\bar{X}_n^c(0), \bar{X}_m^p(0)\right), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r\right)$, the McKean–Vlasov system (14) has a unique solution over any finite time interval $[0, T]$. In addition, this solution depends continuously on the initial condition in the following sense: if $(\bar{X}^1(t), t \in [0, T])$ and $(\bar{X}^2(t), t \in [0, T])$ are two solutions of (14) with two different initial conditions $(\bar{X}^1(0))$ and $(\bar{X}^2(0))$, respectively, then there exists a constant A_T , depending on the time horizon T , such that*

$$\begin{aligned} & \max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\left\| \bar{X}_{n,j}^{1,c} - \bar{X}_{n,j}^{2,c} \right\|_T \right] + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left[\left\| \bar{X}_{n,j}^{1,p} - \bar{X}_{n,j}^{2,p} \right\|_T \right] \\ & \leq \left(\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\left\| \bar{X}_{n,j}^{1,c}(0) - \bar{X}_{n,j}^{2,c}(0) \right\| \right] + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left[\left\| \bar{X}_{n,j}^{1,p}(0) - \bar{X}_{n,j}^{2,p}(0) \right\| \right] \right) e^{A_T}. \end{aligned} \tag{25}$$

Proof. For $1 \leq j \leq r$, with a slight abuse of notation, let

$$e_{j,c} : (x_1^c, x_1^p, \dots, x_r^c, x_r^p) \in (\mathcal{Z}^{2r}) \rightarrow x_j^c \in \mathcal{Z}$$

and

$$e_{j,p} : (x_1^c, x_1^p, \dots, x_r^c, x_r^p) \in (\mathcal{Z}^{2r}) \rightarrow x_j^p \in \mathcal{Z}$$

be the c th and the p th component, respectively, of the j th projection. Moreover, for $t \leq T$, let $p_t : f \in \mathcal{D}([0, T], \mathcal{Z}^{2r}) \rightarrow f(t) \in \mathcal{Z}^{2r}$.

For $M \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$, let $M(t) := M \circ p_t^{-1}$. Consider the system starting at $\bar{X}_0 = (\bar{X}_0^{1,c}, \bar{X}_0^{1,p}, \dots, \bar{X}_0^{r,c}, \bar{X}_0^{r,p})$ and given, at each $t \in (0, T]$, by

$$\begin{aligned} \bar{X}_j^c(t) &= \bar{X}_j^c(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{\bar{X}_j^c(s-) = z'}(z' - z) \mathbb{1}_{\left[0, \lambda_{j,z,z'}^c, v_j^c(s-)\right]}(y) \mathcal{N}_j^c(ds, dy), \\ \bar{X}_j^p(t) &= \bar{X}_j^p(0) + \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{\bar{X}_j^p(s-) = z'}(z' - z) \mathbb{1}_{\left[0, \lambda_{j,z,z'}^p, v_j^p(s-)\right]}(y) \mathcal{N}_j^p(ds, dy), \end{aligned} \tag{26}$$

for $1 \leq j \leq r$, where $\mu_j^c(t) = M(t) \circ e_{j,c}^{-1}$ and $\mu_j^p(t) = M(t) \circ e_{j,p}^{-1}$, the vectors $v_j^c(t)$ and $v_j^p(t)$ are given by (15), and $\{\mathcal{N}_j^c, 1 \leq j \leq r\}$ and $\{\mathcal{N}_j^p, 1 \leq j \leq r\}$ are collections of Poisson random measures on \mathbb{R}^2 whose intensity measures are Lebesgue on \mathbb{R}_+^2 . Denote by ψ and ϕ the mappings that associate to M the solution of this system and its corresponding law. Thus, $\psi(M) = (\bar{X}_j^c, \bar{X}_j^p, 1 \leq j \leq r)$ and $\phi(M) = \mathcal{L}(\bar{X}_j^c, \bar{X}_j^p, 1 \leq j \leq r)$. Observe that if \bar{X} is a solution of (14), then its law is a fixed point of ϕ . Conversely, if M is a fixed point of ϕ for the system (26), then the corresponding solution $\psi(M)$ defines a solution of the limiting system (14). The idea is then to prove the existence of a fixed point of ϕ .

Take $M_1, M_2 \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$. Set $\bar{X}_1 := (\bar{X}_1^{1,c}, \bar{X}_1^{1,p}, \dots, \bar{X}_r^{1,c}, \bar{X}_r^{1,p}) = \psi(M_1)$ and $\bar{X}_2 := (\bar{X}_1^{2,c}, \bar{X}_1^{2,p}, \dots, \bar{X}_r^{2,c}, \bar{X}_r^{2,p}) = \psi(M_2)$. Thus, $\mathcal{L}(\bar{X}_1) = \phi(M_1)$ and $\mathcal{L}(\bar{X}_2) = \phi(M_2)$. Moreover, for all $t \in [0, T]$, define $\mu_1(t) := (\mu_1^{1,c}(t), \mu_1^{1,p}(t), \dots, \mu_r^{1,c}(t), \mu_r^{1,p}(t))$ and $\mu_2(t) := (\mu_1^{2,c}(t), \mu_1^{2,p}(t), \dots, \mu_r^{2,c}(t), \mu_r^{2,p}(t))$ with $\mu_j^{1,c}(t) := M_1(t) \circ e_{j,c}^{-1}$, $\mu_j^{1,p}(t) := M_1(t) \circ e_{j,p}^{-1}$, $\mu_j^{2,c}(t) := M_2(t) \circ e_{j,c}^{-1}$ and $\mu_j^{2,p}(t) := M_2(t) \circ e_{j,p}^{-1}$ for $1 \leq j \leq r$. According to (15), we introduce the following notation:

$$\begin{aligned} v_j^{1,c}(t) &:= (\mu_j^{1,c}(t), \mu_j^{1,p}(t), p_j^c, p_j^p), \\ v_j^{2,c}(t) &:= (\mu_j^{2,c}(t), \mu_j^{2,p}(t), p_j^c, p_j^p), \\ v_{n,j}^{1,p}(t) &:= (\mu_j^{1,c}(t), \mu_1^{1,p}(t), \dots, \mu_r^{1,p}(t), \alpha_j^c, q_{j,1}, \dots, q_{j,r}), \\ v_{n,j}^{2,p}(t) &:= (\mu_j^{2,c}(t), \mu_1^{2,p}(t), \dots, \mu_r^{2,p}(t), \alpha_j^c, q_{j,1}, \dots, q_{j,r}). \end{aligned} \tag{27}$$

We first prove that ϕ is a contraction mapping on $\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$; that is, for any $t \in [0, T]$,

$$\mathcal{W}_{1,t}(\phi(M_1), \phi(M_2)) \leq C(t) \mathbb{E} \left[\int_{[0,t]} \mathcal{W}_{1,s}(M_1, M_2) ds \right]. \tag{28}$$

To this end, for ease of reading, let us introduce the following notation:

$$\Delta_j^c(t) := \|\bar{X}_j^{1,c} - \bar{X}_j^{2,c}\|_t, \quad \Delta_j^p(t) := \|\bar{X}_j^{1,p} - \bar{X}_j^{2,p}\|_t. \tag{29}$$

Indeed, for any $1 \leq j \leq r$ we have that

$$\begin{aligned} \Delta_j^c(t) &\leq \int_{[0,t] \times \mathbb{R}^+} \left| \sum_{(z,z') \in \mathcal{E}} (z' - z) \left\{ \mathbb{1}_{\bar{X}_j^{1,c}(s-)=z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c]}(v_j^{1,c}(s)) \right. \right. \\ &\quad \left. \left. - \mathbb{1}_{\bar{X}_j^{2,c}(s-)=z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c]}(v_j^{2,c}(s)) \right\} \right| \mathcal{N}_j^c(ds, dy). \end{aligned} \tag{30}$$

Using a martingale argument (see (63)) and taking the expectation, by adding and subtracting terms (see (65)) one gets, for any $t \in [0, T]$,

$$\begin{aligned} \mathbb{E} \left[\Delta_j^c(t) \right] &\leq K \mathbb{E} \left[\int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \left| \left(\mathbb{1}_{\bar{X}_j^{1,c}(s)=z} - \mathbb{1}_{\bar{X}_j^{2,c}(s)=z} \right) \lambda_{j,z,z'}^c(v_j^{1,c}(s)) \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\bar{X}_j^{2,c}(s)=z} \left(\lambda_{j,z,z'}^c(v_j^{1,c}(s)) - \lambda_{j,z,z'}^c(v_j^{2,c}(s)) \right) \right| ds \right]. \end{aligned} \tag{31}$$

Recall the definition of the functions $\lambda_{j,z,z'}^c$ in (16). Given that $\mu_j^c(t)$ and $\mu_j^p(t)$ are probability measures and using the boundedness of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$, one easily gets that

$$\lambda_{j,z,z'}^c(v_j^{1,c}(s)) \leq \bar{\gamma}, \tag{32}$$

and

$$\begin{aligned} \left| \lambda_{j,z,z'}^c(v_j^{1,c}(s)) - \lambda_{j,z,z'}^c(v_j^{2,c}(s)) \right| &\leq p_j^c \left| \gamma_{z,z'}^c(\mu_j^{1,c}(s) - \mu_j^{2,c}(s)) \right| + p_j^p \left| \gamma_{z,z'}^p(\mu_j^{1,p}(s) - \mu_j^{2,p}(s)) \right| \\ &\leq p_j^c \bar{\gamma} d_{BL}(\mu_j^{1,c}(s), \mu_j^{2,c}(s)) + p_j^p \bar{\gamma} d_{BL}(\mu_j^{1,p}(s), \mu_j^{2,p}(s)). \end{aligned} \tag{33}$$

Therefore one obtains

$$\begin{aligned} \mathbb{E}[\Delta_j^c(t)] &\leq K|\mathcal{E}|\bar{\gamma}\mathbb{E}^0 \left[\int_{[0,t]} \left(|\bar{X}_j^{1,c}(s) - \bar{X}_j^{2,c}(s)| + p_j^c d_{BL}(\mu_j^{1,c}(s), \mu_j^{2,c}(s)) \right. \right. \\ &\quad \left. \left. + p_j^p d_{BL}(\mu_j^{1,p}(s), \mu_j^{2,p}(s)) \right) ds \right]. \end{aligned} \tag{34}$$

Using (17) and the same steps as previously, one finds, for any $1 \leq j \leq r$,

$$\begin{aligned} \mathbb{E}[\Delta_j^p(t)] &\leq K|\mathcal{E}|\bar{\gamma}\mathbb{E} \left[\int_{[0,t]} \left(|\bar{X}_j^{1,p}(s) - \bar{X}_j^{2,p}(s)| + \alpha_j^c d_{BL}(\mu_j^{1,c}(s), \mu_j^{2,c}(s)) \right. \right. \\ &\quad \left. \left. + q_{j1} d_{BL}(\mu_1^{1,p}(s), \mu_1^{2,p}(s)) + \dots + q_{jr} d_{BL}(\mu_r^{1,p}(s), \mu_r^{2,p}(s)) \right) ds \right]. \end{aligned} \tag{35}$$

On the one hand, from the Kantorovich–Rubinstein theorem, one has that for $1 \leq j \leq r$ and $\alpha \in \{c, p\}$,

$$\mathcal{W}_1(\mu_j^{1,\alpha}(s), \mu_j^{2,\alpha}(s)) = \sup_{\|g\|_{L^1} \leq 1} \left| \langle \mu_j^{1,\alpha}(s), g \rangle - \langle \mu_j^{2,\alpha}(s), g \rangle \right|, \tag{36}$$

where the supremum is taken over the functions g with Lipschitz constant 1. Therefore,

$$d_{BL}(\mu_j^{1,\alpha}(s), \mu_j^{2,\alpha}(s)) \leq \mathcal{W}_1(\mu_j^{1,\alpha}(s), \mu_j^{2,\alpha}(s)). \tag{37}$$

On the other hand, one can easily verify that

$$\mathcal{W}_1(\mu_j^{1,\alpha}(s), \mu_j^{2,\alpha}(s)) \leq \mathcal{W}_{1,s}(M_1, M_2). \tag{38}$$

Thus, using (37) and (38) and taking the supremum over $1 \leq j \leq r$ in (34) and (35), one obtains

$$\mathbb{E} \left[\sup_{1 \leq j \leq r} \Delta_j^c(t) \right] \leq K|\mathcal{E}|\bar{\gamma}\mathbb{E} \left[\int_{[0,t]} \left(\sup_{1 \leq j \leq r} \Delta_j^c(s) + \mathcal{W}_{1,s}(M_1, M_2) \right) ds \right] \tag{39}$$

and

$$\mathbb{E} \left[\sup_{1 \leq j \leq r} \Delta_j^p(t) \right] \leq K|\mathcal{E}|\bar{\gamma} \mathbb{E} \left[\int_{[0,t]} \left(\sup_{1 \leq j \leq r} \Delta_j^p(s) + \mathcal{W}_{1,s}(M_1, M_2) \right) ds \right]. \tag{40}$$

Adding the two last inequalities side by side and applying Grönwall’s lemma leads to

$$\mathbb{E} \left[\sup_{1 \leq j \leq r} \Delta_j^c(t) + \sup_{1 \leq j \leq r} \Delta_j^p(t) \right] \leq K|\mathcal{E}|\bar{\gamma} \mathbb{E} \left[\int_{[0,t]} \mathcal{W}_{1,s}(M_1, M_2) ds \right] e^{K|\mathcal{E}|\bar{\gamma}t}. \tag{41}$$

Hence,

$$\mathbb{E}[\|\bar{X}^1 - \bar{X}^2\|_t] \leq C(t) \mathbb{E} \left[\int_{[0,t]} \mathcal{W}_{1,s}(M_1, M_2) ds \right], \tag{42}$$

with $C(t) = K|\mathcal{E}|\bar{\gamma}e^{K|\mathcal{E}|\bar{\gamma}t}$. From the definition of the Wasserstein distance, it is easy to observe that

$$\mathcal{W}_{1,t}(\phi(M_1), \phi(M_2)) \leq \mathbb{E}[\|\bar{X}^1 - \bar{X}^2\|_t],$$

from which one deduces (28).

Consider now the following recursive scheme:

- $M_0 \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$;
- $M_{k+1} = \phi(M_k), \quad k \geq 0$.

By iterating the formula in (28) and using the fact that $\mathcal{W}_{1,t}(M_1, M_0)$ is increasing in t , one finds that for all $k \geq 0$,

$$\mathcal{W}_{1,t}(M_{k+2}, M_{k+1}) \leq \frac{(tC(t))^k}{k!} \mathcal{W}_{1,t}(M_1, M_2).$$

Moreover, it is easy to verify that $\mathbb{E}[\|\bar{X}\|_T] < \infty$, where $\bar{X} = \psi(M)$ for any $M \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$, from which we deduce that $\mathcal{W}_{1,t}(M_1, M_2) < \infty$ and thus the sequence $\{M_k\}_{k \geq 0}$ is a Cauchy sequence. Note that the space $\mathcal{P}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$ endowed with the Wasserstein distance $\mathcal{W}_{p,T}$ is complete (see [11]). Hence the sequence $\{M_k\}_{k \geq 0}$ converges to some measure M in $\mathcal{P}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$ which is a fixed point of ϕ on $\mathcal{P}_1(\mathcal{D}([0, T], \mathcal{Z}^{2r}))$. This proves the existence of the solution of the equation in (26) and thus that of the equation in (14). Uniqueness follows from again using (28) and Grönwall’s lemma.

Let $(\bar{X}^1(t)) := (\bar{X}_{n,j}^{1,c}(t), \bar{X}_{m,j}^{1,p}(t), n \in C_j^c, m \in C_j^p, 1 \leq j \leq r)$ and $(\bar{X}^2(t)) := (\bar{X}_{n,j}^{2,c}(t), \bar{X}_{m,j}^{2,p}(t), n \in C_j^c, m \in C_j^p, 1 \leq j \leq r)$ be two solutions of (14) with respective initial conditions $(\bar{X}^1(0))$ and $(\bar{X}^2(0))$. Denote by $\mu_j^{1,c}(t) := \mathcal{L}(\bar{X}_{n,j}^{1,c}(t))$ and $\mu_j^{1,p}(t) := \mathcal{L}(\bar{X}_{m,j}^{1,p}(t))$, for $1 \leq j \leq r$, the probability measures corresponding to the first solution. Similarly, denote by $\mu_j^{2,c}(t) := \mathcal{L}(\bar{X}_{n,j}^{2,c}(t))$ and $\mu_j^{2,p}(t) := \mathcal{L}(\bar{X}_{m,j}^{2,p}(t))$ the probability measures corresponding to the second solution. Again, for ease of reading, let us introduce the following notation:

$$\Delta_{n,j}^c(t) := \|\bar{X}_{n,j}^{1,c} - \bar{X}_{n,j}^{2,c}\|_t, \quad \Delta_{n,j}^p(t) := \|\bar{X}_{n,j}^{1,p} - \bar{X}_{n,j}^{2,p}\|_t. \tag{43}$$

Using this together with the notation in (27), one finds that for any $n \in C_j^c$,

$$\begin{aligned} \Delta_{n,j}^c(t) &\leq \left| \bar{X}_{n,j}^{1,c}(0) - \bar{X}_{n,j}^{2,c}(0) \right| \\ &\quad + \int_{[0,t] \times \mathbb{R}_+} \left| \sum_{(z,z') \in \mathcal{E}} (z' - z) \left\{ \mathbb{1}_{\bar{X}_{n,j}^{1,c}(s-) = z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_j^{1,c}(s))]}(y) \right. \right. \\ &\quad \left. \left. - \mathbb{1}_{\bar{X}_{n,j}^{2,c}(s-) = z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_j^{2,c}(s))]}(y) \right\} \right| \mathcal{N}_{n,j}^c(ds, dy). \end{aligned} \tag{44}$$

Using a martingale argument (see (63)), taking the conditional expectation E^0 given $(\bar{X}^1(0), \bar{X}^2(0))$, and finally adding and subtracting terms (see (65)), one gets that for $t \in [0, T]$,

$$\begin{aligned} \mathbb{E}^0 \left[\Delta_{n,j}^c(t) \right] &\leq \left| \bar{X}_{n,j}^{1,c}(0) - \bar{X}_{n,j}^{2,c}(0) \right| \\ &\quad + K \mathbb{E}^0 \left[\int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \left| \left(\mathbb{1}_{\bar{X}_{n,j}^{1,c}(s) = z} - \mathbb{1}_{\bar{X}_{n,j}^{2,c}(s) = z} \right) \lambda_{j,z,z'}^c(v_j^{1,c}(s)) \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\bar{X}_{n,j}^{2,c}(s-) = z} \left(\lambda_{j,z,z'}^c(v_j^{1,c}(s)) - \lambda_{j,z,z'}^c(v_j^{2,c}(s)) \right) \right| ds \right]. \end{aligned} \tag{45}$$

Since $\mu_j^{1,c}(t)$ and $\mu_j^{1,p}(t)$ are probability measures, it is easy to see using (16) and the boundedness of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ that

$$\lambda_{j,z,z'}^c(v_j^{1,c}(s)) \leq \bar{\gamma}. \tag{46}$$

Additionally, the Lipschitz property of the functions $\gamma_{z,z'}^{j,c}$ leads to

$$\left| \lambda_{j,z,z'}^c(v_j^{1,c}(s)) - \lambda_{j,z,z'}^c(v_j^{2,c}(s)) \right| \leq p_j^c \bar{\gamma} \mathbb{E} \left[\left| \bar{X}_{n,j}^{1,c}(s) - \bar{X}_{n,j}^{2,c}(s) \right| \right] + p_j^p \bar{\gamma} \mathbb{E} \left[\left| \bar{X}_{1,j}^{1,p}(s) - \bar{X}_{1,j}^{2,p}(s) \right| \right]. \tag{47}$$

Therefore one obtains

$$\mathbb{E}^0 \left[\Delta_{n,j}^c(t) \right] \leq \Delta_{n,j}^c(0) + K \bar{\gamma} |\mathcal{E}| \int_{[0,t]} \left(\mathbb{E}^0 \left[\Delta_{n,j}^c(s) \right] + p_j^c \mathbb{E} \left[\Delta_{n,j}^c(s) \right] + p_j^p \mathbb{E} \left[\Delta_{1,j}^p(s) \right] \right) ds. \tag{48}$$

Taking the expectation on both sides of the last inequality, and recalling that $p_j^c + p_j^p = 1$ for all $1 \leq j \leq r$, one gets

$$\mathbb{E} \left[\Delta_{n,j}^c(t) \right] \leq \mathbb{E} \left[\Delta_{n,j}^c(0) \right] + 2K \bar{\gamma} |\mathcal{E}| \int_{[0,t]} \left(\mathbb{E} \left[\Delta_{n,j}^c(s) \right] + \mathbb{E} \left[\Delta_{1,j}^p(s) \right] \right) ds. \tag{49}$$

Thus,

$$\mathbb{E} \left[\Delta_{n,j}^c(t) \right] \leq \max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E} \left[\Delta_{m,j}^c(0) \right] + 2K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E} \left[\Delta_{m,j}^c(s) \right] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\Delta_{m,j}^p(s) \right] \right) ds. \tag{50}$$

Taking the maximum over $n \in C_j^c$ and $1 \leq j \leq r$ gives

$$\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\Delta_{n,j}^c(t) \right] \leq \max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E} \left[\Delta_{m,j}^c(0) \right] + 2K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E} \left[\Delta_{m,j}^c(s) \right] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\Delta_{m,j}^p(s) \right] \right) ds. \tag{51}$$

Using similar arguments, one finds that for any $1 \leq j \leq r$ and $n \in C_j^p$,

$$\mathbb{E}^0 \left[\Delta_{n,j}^p(t) \right] \leq |\bar{X}_{n,j}^{1,p}(0) - \bar{X}_{n,j}^{2,p}(0)| + K\mathbb{E}^0 \left[\int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \left| \left(\bar{X}_{n,j}^{1,p}(s) - \bar{X}_{n,j}^{2,p}(s) \right) \lambda_{j,z,z'}^p \left(v_j^{1,p}(s) \right) + \left(\lambda_{j,z,z'}^p \left(v_j^{1,p}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^{2,p}(s) \right) \right) \right| \right] ds. \tag{52}$$

By (17) and the Lipschitz boundedness property of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ one finds that

$$\mathbb{E}^0 \left[\Delta_{n,j}^p(t) \right] \leq \Delta_{n,j}^p(0) + K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\mathbb{E}^0 \left[\Delta_{n,j}^p(s) \right] + \alpha_j^c \mathbb{E} \left[\Delta_{1,j}^c(s) \right] + q_{j1} \mathbb{E} \left[\Delta_{1,1}^p(s) \right] + \dots + q_{jr} \mathbb{E} \left[\Delta_{1,r}^p(s) \right] \right) ds. \tag{53}$$

Taking the expectation on both sides of the last inequality gives

$$\mathbb{E} \left[\Delta_{n,j}^p(t) \right] \leq \mathbb{E} \left[\Delta_{n,j}^p(0) \right] + K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\mathbb{E} \left[\Delta_{n,j}^p(s) \right] + \alpha_j^c \mathbb{E} \left[\Delta_{1,j}^c(s) \right] + q_{j1} \mathbb{E} \left[\Delta_{1,1}^p(s) \right] + \dots + q_{jr} \mathbb{E} \left[\Delta_{1,r}^p(s) \right] \right) ds. \tag{54}$$

Recall that, for all $1 \leq j \leq r$, $\alpha_j^c + q_{j1} + \dots + q_{jr} = 1$. Therefore

$$\mathbb{E} \left[\Delta_{n,j}^p(t) \right] \leq \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\Delta_{m,j}^p(0) \right] + K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E} \left[\Delta_{m,j}^c(s) \right] + (r+1) \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\Delta_{m,j}^p(s) \right] \right) ds. \tag{55}$$

Taking the maximum over $n \in C_j^p$ and $1 \leq j \leq r$ gives

$$\begin{aligned} \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E}[\Delta_{n,j}^p(t)] &\leq \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E}[\Delta_{m,j}^p(0)] \\ &\quad + K\bar{\gamma}|\mathcal{E}| \int_{[0,t]} \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E}[\Delta_{m,j}^c(s)] + (r+1) \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E}[\Delta_{m,j}^p(s)] \right) ds. \end{aligned} \tag{56}$$

Now (51) and (56) together lead to

$$\begin{aligned} &\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E}[\Delta_{n,j}^c(t)] + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E}[\Delta_{n,j}^p(t)] \\ &\leq \max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E}[\Delta_{m,j}^c(0)] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E}[\Delta_{m,j}^p(0)] \\ &\quad + 4K\bar{\gamma}|\mathcal{E}|(r+2) \int_{[0,t]} \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E}[\Delta_{m,j}^c(s)] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E}[\Delta_{m,j}^p(s)] \right) ds. \end{aligned} \tag{57}$$

Then Grönwall’s lemma gives

$$\begin{aligned} &\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E}[\Delta_{n,j}^c(t)] + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E}[\Delta_{n,j}^p(t)] \\ &\leq \left(\max_{1 \leq j \leq r} \max_{m \in C_j^c} \mathbb{E}[\Delta_{m,j}^c(0)] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E}[\Delta_{m,j}^p(0)] \right) \exp \{t4K\bar{\gamma}|\mathcal{E}|(r+2)\}. \end{aligned} \tag{58}$$

Defining $A_t = 8K\bar{\gamma}|\mathcal{E}|(2+r)t$ leads to (25). The theorem is proved. □

5. Laws of large numbers and propagation of chaos

This section investigates the weak convergence of the finite particle system represented by the stochastic differential equation in (12) towards the limiting McKean–Vlasov system (14) as the total number of particles N tends to infinity. In particular, as the main results of this section, we establish propagation of chaos in Theorem 5.1 and laws of large numbers in Corollary 5.1.

Let us start this section by recalling the notions of multi-exchangeability and multi-chaoticity introduced in [40].

Definition 5.1. A sequence of random variables $(X_{n,k}, 1 \leq n \leq N_k, 1 \leq k \leq K)$ indexed by $(N_k, 1 \leq k \leq K) \in \mathbb{N}^K$ is said to be multi-exchangeable if its law is invariant under permutation of the indexes within the classes; that is, for any permutations σ_k of $\{1, \dots, N_k\}$ for $1 \leq k \leq K$, the following equality holds in distribution:

$$(X_{\sigma_k(n),k}, 1 \leq n \leq N_k, 1 \leq k \leq K) \stackrel{dist}{=} (X_{n,k}, 1 \leq n \leq N_k, 1 \leq k \leq K).$$

A sequence of random variables $(X_{n,k}, 1 \leq n \leq N_k, 1 \leq k \leq K)$ indexed by $(N_k, 1 \leq k \leq K) \in \mathbb{N}^K$ is $P_1 \otimes \dots \otimes P_K$ -multi-chaotic if, for any $m \geq 1$, the convergence in distribution

$$\lim_{N \rightarrow \infty} (X_{n,k}, 1 \leq n \leq m, 1 \leq k \leq K) \stackrel{dist}{=} P_1^{\otimes m} \otimes \dots \otimes P_K^{\otimes m}$$

holds for the topology of uniform convergence on compact sets, where P_k , for $1 \leq k \leq K$, is a probability distribution on \mathbb{R}_+ , and with the convention that N goes to infinity when $\min N_k$ goes to infinity.

5.1. Propagation of chaos

The following result establishes the weak convergence of the pre-limiting system (12) towards the McKean–Vlasov system (14) as the number of particles N goes to infinity, and thus its multi-chaoticity.

Theorem 5.1. *Suppose that Condition 4.1 holds true. Moreover, suppose that the initial conditions $(X_{n,j}^c(0), X_{m,j}^p(0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$ are multi-exchangeable and $v^{1,c} \otimes v^{1,p} \otimes \dots \otimes v^{r,c} \otimes v^{r,p}$ -multi-chaotic. Then, for any $t \in [0, T]$, as $N \rightarrow \infty$,*

$$\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\|X_{n,j}^c - \bar{X}_{n,j}^c\|_t \right] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\|X_{m,j}^p - \bar{X}_{m,j}^p\|_t \right] \rightarrow 0, \tag{59}$$

and the sequence of processes $\left((X_n^{c,N}(t), X_m^{p,N}(t), t \geq 0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r) \right)$, the solutions of the stochastic differential equation (12) with initial conditions $(X_n^{c,N}(0), X_m^{p,N}(0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$, is $P_{\bar{X}}$ -multi-chaotic, where $P_{\bar{X}} = \mu_1^c \otimes \mu_1^p \otimes \dots \otimes \mu_r^c \otimes \mu_r^p$ is the distribution of the process $\left((\bar{X}_n^c(t), \bar{X}_m^p(t), t \geq 0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r) \right)$, the solution of the limiting stochastic differential equation (14) with initial distribution $v^{1,c} \otimes v^{1,p} \otimes \dots \otimes v^{r,c} \otimes v^{r,p}$.

Before proceeding to the proof, we recall, without proof, an elementary result on (conditionally) independent and identically distributed (i.i.d.) random variables.

Lemma 5.1. *Let $\{S_i : i = 1, \dots, n\}$ be a collection of \mathbb{S} -valued random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where \mathbb{S} is a Polish space. Suppose that S_1, \dots, S_n are conditionally i.i.d. given some σ -algebra $\mathcal{G} \subset \mathcal{F}$. Then, for any $k \in \mathbb{N}$, there exists a positive and finite constant $0 < a_k < \infty$ such that*

$$\sup_{\|f\|_\infty \leq 1} \mathbb{E} \left| \frac{1}{n} \sum_{i=1}^n (f(S_i) - \mathbb{E}[f(S_i)|\mathcal{G}]) \right|^k \leq \frac{a_k}{n^{k/2}}. \tag{60}$$

Proof of Theorem 5.1. We use a coupling method. Let $X(t) = \left((X_{n,j}^c(t), X_{m,j}^p(t)), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r \right)$ be the solution of the stochastic differential equation (12) with initial conditions $X(0) = \left((X_{n,j}^c(0), X_{m,j}^p(0)), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r \right)$. Moreover, let $Y(t) = \left((Y_{n,j}^c(t), Y_{m,j}^p(t)), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r \right)$ be the solution of the limiting stochastic differential equation (14) with the same initial conditions as $X(t)$, i.e., $Y(0) = X(0)$. Also, let the processes $Y(t)$ and $X(t)$ be defined on the same probability space by taking the same sequences of Poisson random measures $\{\mathcal{N}_{n,j}^c\}$ and $\{\mathcal{N}_{n,j}^p\}$ in both cases. We first prove that these two processes are asymptotically close, that is, for any $t \in [0, T]$,

$$\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] + \max_{1 \leq j \leq r} \max_{m \in C_j^p} \mathbb{E} \left[\|X_{m,j}^p - Y_{m,j}^p\|_t \right] \rightarrow 0, \text{ as } N \rightarrow \infty. \tag{61}$$

To this end, we treat the central and peripheral nodes in two separate steps. For convenience, define

$$o(n) := \frac{1}{deg(n) + 1}.$$

Step 1. Fix $1 \leq j \leq r$. For each central node $n \in C_j^c$ and any $t \in [0, T]$,

$$\begin{aligned} \mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] &= \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X_{n,j}^c(s) - Y_{n,j}^c(s) \right| \right] \\ &\leq \mathbb{E} \left[\left| \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{X_{n,j}^c(s-) = z'(z-z)} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s))]}(y) \mathcal{N}_{n,j}^c(ds, dy) \right. \right. \\ &\quad \left. \left. - \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{Y_{n,j}^c(s-) = z'(z-z)} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_j^c(s))]}(y) \mathcal{N}_{n,j}^c(ds, dy) \right| \right] \\ &\leq \mathbb{E} \left[\int_{[0,t] \times \mathbb{R}_+} \left| \sum_{(z,z') \in \mathcal{E}} (z' - z) \left\{ \mathbb{1}_{X_{n,j}^c(s-) = z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s))]}(y) \right. \right. \right. \\ &\quad \left. \left. \left. - \mathbb{1}_{Y_{n,j}^c(s-) = z} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_j^c(s))]}(y) \right\} \mathcal{N}_{n,j}^c(ds, dy) \right| \right]. \end{aligned} \tag{62}$$

Denote by \mathcal{F}_t the filtration generated by the Poisson random measures and defined by

$$\mathcal{F}_t := \sigma \left\langle \mathcal{N}_{n,j}^c(A \times B) : n \in C_j^c \cup \mathcal{N}_{m,j}^p(A \times B) : m \in C_j^p, A \in \mathcal{B}(\mathbb{R}_+), B \in \mathcal{B}([0, T]) \right\rangle.$$

Then $X_{n,j}^c(t)$ and $Y_{n,j}^c(t)$ are adapted to the filtration \mathcal{F}_t . Therefore, the two processes

$$\begin{aligned} &\int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{X_{n,j}^c(s-) = z'(z-z)} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s))]}(y) \left[\mathcal{N}_{n,j}^c(ds, dy) - dsdy \right], \\ &\int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{Y_{n,j}^c(s-) = z'(z-z)} \mathbb{1}_{[0,\lambda_{j,z,z'}^c(v_j^c(s))]}(y) \left[\mathcal{N}_{n,j}^c(ds, dy) - dsdy \right] \end{aligned} \tag{63}$$

are \mathcal{F}_t -martingales. Furthermore, (62) reduces to

$$\begin{aligned} \mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] &\leq \mathbb{E} \left[\int_{[0,t]} \left| \sum_{(z,z') \in \mathcal{E}} (z' - z) \left\{ \mathbb{1}_{X_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s)) \right. \right. \right. \\ &\quad \left. \left. \left. - \mathbb{1}_{Y_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_j^c(s)) \right\} \right| ds \right]. \end{aligned} \tag{64}$$

Recall that $K = |\mathcal{Z}|$ is the cardinality of the set \mathcal{Z} . By adding and subtracting terms we obtain

$$\begin{aligned} &\mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] \\ &\leq K \mathbb{E} \left[\int_{[0,t]} \left| \sum_{(z,z') \in \mathcal{E}} \left\{ \mathbb{1}_{X_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s)) - \mathbb{1}_{Y_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s)) \right. \right. \right. \\ &\quad \left. \left. \left. + \mathbb{1}_{Y_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_{n,j}^{c,N}(s)) - \mathbb{1}_{Y_{n,j}^c(s) = z} \lambda_{j,z,z'}^c(v_j^c(s)) \right\} \right| ds \right] \end{aligned}$$

$$\begin{aligned} &\leq K \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \mathbb{E} \left[\left| \left(\mathbb{1}_{X_{n,j}^c(s)=z} - \mathbb{1}_{Y_{n,j}^c(s)=z} \right) \lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{Y_{n,j}^c(s)=z} \left(\lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) - \lambda_{j,z,z'}^c \left(v_j^c(s) \right) \right) \right| \right] ds. \end{aligned} \tag{65}$$

The goal now is to bound the right-hand side of (65). Let us start with the second term. Again, by adding and subtracting terms we get

$$\begin{aligned} &\mathbb{E} \left[\left| \mathbb{1}_{Y_{n,j}^c(s)=z} \left(\lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) - \lambda_{j,z,z'}^c \left(v_j^c(s) \right) \right) \right| \right] \\ &\leq \mathbb{E} \left[\left| \lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) - \lambda_{j,z,z'}^c \left(v_j^c(s) \right) \right| \right] \\ &= \mathbb{E} \left[\left| \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,j}^p(s) \right) \right) \right. \right. \\ &\quad - \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(Y_{m,j}^p(s) \right) \right) \\ &\quad + \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(Y_{m,j}^p(s) \right) \right) \\ &\quad \left. - \left(p_j^c \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds + p_j^p \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_j^p(s) ds \right) \right| \right] \\ &\leq \mathbb{E} \left[\left| \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,j}^p(s) \right) \right) \right. \right. \\ &\quad \left. \left. - \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(Y_{m,j}^p(s) \right) \right) \right| \right] \\ &+ \mathbb{E} \left[\left| \left(o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(Y_{m,j}^p(s) \right) \right) \right. \right. \\ &\quad \left. \left. - \left(p_j^c \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds + p_j^p \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_j^p(s) ds \right) \right| \right]. \end{aligned} \tag{66}$$

Denote by \mathcal{J}_1 and \mathcal{J}_2 respectively the first and the second expectation in the right-hand side of (66). Then, from the Lipschitz property of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$, \mathcal{J}_1 is bounded as follows:

$$\begin{aligned} \mathcal{J}_1 &\leq o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} L_{\gamma} \mathbb{E} \left[\left| X_{m,j}^c(s) - Y_{m,j}^c(s) \right| \right] + o(n) \sum_{m \in \mathfrak{N}_j^p(n)} L_{\gamma} \mathbb{E} \left[\left| X_{m,j}^p(s) - Y_{m,j}^p(s) \right| \right] \\ &\leq \varrho_{n,j}^c L_{\gamma} \max_{m \in C_j^c} \mathbb{E} \left\| X_{m,j}^c - Y_{m,j}^c \right\|_s + \varrho_{n,j}^p L_{\gamma} \max_{m \in C_j^p} \mathbb{E} \left\| X_{m,j}^p - Y_{m,j}^p \right\|_s, \end{aligned} \tag{67}$$

where L_γ is the maximum Lipschitz constant of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ for all $(z, z') \in \mathcal{Z}^2$. Moreover, by adding and subtracting terms and using the fact that both $\{Y_{n,j}^c(s)\}$ and $\{Y_{n,j}^p(s)\}$ are sequences of i.i.d. random variables, \mathcal{J}_2 can be bounded as follows:

$$\begin{aligned} \mathcal{J}_2 &\leq \mathbb{E} \left[\left| o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c}(Y_{m,j}^c(s)) - p_j^c \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds \right| \right] \\ &\quad + \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) - p_j^p \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_j^p(s) ds \right| \right] \\ &\leq \mathbb{E} \left[\left| p_j^c \frac{1}{M_j^c(n) + 1} \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \left(\gamma_{z,z'}^{j,c}(Y_{m,j}^c(s)) - \mathbb{E} \left[\gamma_{z,z'}^{j,c}(Y_{m,j}^c(s)) \right] \right) \right| \right] \\ &\quad + \left| o(n) - \frac{p_j^c}{M_j^c(n) + 1} \right| \mathbb{E} \left[\left| \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c}(Y_{m,j}^c(s)) \right| \right] \\ &\quad + \mathbb{E} \left[\left| p_j^p \frac{1}{M_j^p(n)} \sum_{m \in \mathfrak{N}_j^p(n)} \left(\gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) - \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right] \right) \right| \right] \\ &\quad + \left| o(n) - \frac{p_j^p}{M_j^p(n)} \right| \mathbb{E} \left[\left| \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right| \right]. \end{aligned} \tag{68}$$

Note that, by the exchangeability of $\{Y_{n,j}^c(s), n \in C_j^c\}$ and the boundedness of the functions $\gamma_{z,z'}^{j,c}$, one obtains

$$\left| o(n) - \frac{p_j^c}{M_j^c(n) + 1} \right| \mathbb{E} \left[\left| \sum_{m \in \{n\} \cup \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c}(Y_{m,j}^c(s)) \right| \right] = |q_{n,j}^c - p_j^c| \mathbb{E} \left[\left| \gamma_{z,z'}^{j,c}(Y_{1,j}^c(s)) \right| \right] \leq \bar{\gamma} |q_{n,j}^c - p_j^c|. \tag{69}$$

In the same manner, the fourth term in the right member of (68) is also bounded as follows:

$$\left| o(n) - \frac{p_j^p}{M_j^p(n)} \right| \mathbb{E} \left[\left| \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right| \right] \leq \bar{\gamma} |q_{n,j}^c - p_j^p|. \tag{70}$$

Furthermore, using (60), the first and third expectations in (68) are bounded by $\frac{\kappa_1 p_j^c}{\sqrt{M_j^c(n)+1}}$ and $\frac{\kappa_2 p_j^p}{\sqrt{M_j^p(n)}}$, respectively, where κ_1 and κ_2 are positive constants.

Now let us take a look at the first term of the right-hand side of (65). Since $X_{n,j}^c$ and $Y_{n,j}^c$ are \mathcal{Z} -valued, and \mathcal{Z} is a subset of \mathbb{N} , one easily sees that

$$\begin{aligned} \mathbb{E} \left[\left| \lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) \left(\mathbb{1}_{X_{n,j}^c(s)=z} - \mathbb{1}_{Y_{n,j}^c(s)=z} \right) \right| \right] &\leq \mathbb{E} \left[\left| \lambda_{j,z,z'}^c \left(v_{n,j}^{c,N}(s) \right) \left| X_{n,j}^c(s) - Y_{n,j}^c(s) \right| \right] \\ &\leq \bar{\gamma} \mathbb{E} \left| X_{n,j}^c(s) - Y_{n,j}^c(s) \right| \\ &\leq \bar{\gamma} \mathbb{E} \left\| X_{n,j}^c - Y_{n,j}^c \right\|_s, \end{aligned} \tag{71}$$

where $\bar{\gamma}$ is the upper bound of the functions $\gamma_{z,z'}^{j,c}$ and $\gamma_{z,z'}^{j,p}$ for all $(z, z') \in \mathcal{Z}^2$. Finally, by combining (65), (66), (67), (68), (69), (70), and (71), one obtains

$$\begin{aligned} \mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] &\leq K|\mathcal{E}| \int_0^t \left[\bar{\gamma} \mathbb{E} \|X_{n,j}^c - Y_{n,j}^c\|_s + \varrho_{n,j}^c L_\gamma \max_{m \in C_j^c} \mathbb{E} \|X_{m,j}^c - Y_{m,j}^c\|_s \right. \\ &\quad \left. + \varrho_{n,j}^c L_\gamma \max_{m \in C_j^p} \mathbb{E} \|X_{m,j}^p - Y_{m,j}^p\|_s + \bar{\gamma} \left| \varrho_{n,j}^c - p_j^c \right| \right. \\ &\quad \left. + \bar{\gamma} \left| \varrho_{n,j}^c - p_j^p \right| + \frac{\kappa_1 p_j^c}{\sqrt{M_j^c(n) + 1}} + \frac{\kappa_2 p_j^p}{\sqrt{M_j^p(n)}} \right] ds, \end{aligned} \tag{72}$$

where $|\mathcal{E}|$ stands for the cardinality of the set of edges \mathcal{E} of the graph $(\mathcal{Z}, \mathcal{E})$. Recall that $\text{deg}(n) = M_j^c(n) + M_j^p(n)$ for any $n \in C_j^c$. Using this and then taking the maximum over $n \in C_j^c$ in (72) and over $1 \leq j \leq r$, one finally obtains

$$\begin{aligned} &\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\|X_{n,j}^c - Y_{n,j}^c\|_t \right] \\ &\leq K|\mathcal{E}| \int_0^t \left[(\bar{\gamma} + L_\gamma) \max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \|X_{n,j}^c - Y_{n,j}^c\|_s + L_\gamma \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \|X_{n,j}^p - Y_{n,j}^p\|_s \right. \\ &\quad \left. + \sum_{1 \leq j \leq r} \left(\bar{\gamma} \max_{n \in C_j^c} \left| \varrho_{n,j}^c - p_j^c \right| + \bar{\gamma} \max_{n \in C_j^c} \left| \varrho_{n,j}^c - p_j^p \right| + \max_{n \in C_j^c} \frac{\kappa_1 p_j^c}{\sqrt{M_j^c(n) + 1}} + \max_{n \in C_j^c} \frac{\kappa_2 p_j^p}{\sqrt{M_j^p(n)}} \right) \right] ds. \end{aligned} \tag{73}$$

Step 2. Fix a block $1 \leq j \leq r$ and a peripheral node $n \in C_j^p$. For any $t \in [0, T]$, one has

$$\begin{aligned} \mathbb{E} \left[\|X_{n,j}^p - Y_{n,j}^p\|_t \right] &= \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X_{n,j}^p(s) - Y_{n,j}^p(s) \right| \right] \\ &\leq \mathbb{E} \left[\left| \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{X_{n,j}^p(s)=z} (z' - z) \mathbb{1}_{[0,\lambda_{j,z,z'}^p(v_{n,j}^{p,N}(s))]}(y) \mathcal{N}_{n,j}^p(ds, dy) \right. \right. \\ &\quad \left. \left. - \int_{[0,t] \times \mathbb{R}_+} \sum_{(z,z') \in \mathcal{E}} \mathbb{1}_{Y_{n,j}^p(s)=z} (z' - z) \mathbb{1}_{[0,\lambda_{j,z,z'}^p(v_j^p(s))]}(y) \mathcal{N}_{n,j}^p(ds, dy) \right| \right] \\ &\leq K \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \mathbb{E} \left[\left| \left(\mathbb{1}_{X_{n,j}^p(s)=z} - \mathbb{1}_{Y_{n,j}^p(s)=z} \right) \lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{Y_{n,j}^p(s)=z} \left(\lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^p(s) \right) \right) \right| \right] ds, \end{aligned} \tag{74}$$

where the last inequality is obtained by following the same steps as in (63) and (65). Again, given that $X_{n,j}^p$ and $Y_{n,j}^p$ are \mathcal{Z} -valued and that $\mathcal{Z} \subset \mathbb{N}$, the first expectation in the right-hand side of (74) can be bounded as follows:

$$\mathbb{E} \left[\left| \left(\mathbb{1}_{X_{n,j}^p(s)=z} - \mathbb{1}_{Y_{n,j}^p(s)=z} \right) \lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) \right| \right] \leq \bar{\gamma} \mathbb{E} \left[\|X_{n,j}^p - Y_{n,j}^p\|_s \right]. \tag{75}$$

It remains to bound the second term in the right-hand side of (74). Using Condition 4.1 one gets

$$\begin{aligned}
 & \mathbb{E} \left[\left| \mathbb{1}_{Y_{n,j}^p(s)=z} \left(\lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^p(s) \right) \right) \right| \right] \\
 & \leq \mathbb{E} \left[\left| \lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^p(s) \right) \right| \right] \\
 & = \mathbb{E} \left[\left| o(n) \left(\sum_{m \in \mathfrak{N}_1^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) + \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,1}^p(s) \right) + \dots + \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,j}^p(s) \right) + \dots \right. \right. \right. \\
 & \qquad \qquad \qquad \left. \left. \dots + \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,r}^p(s) \right) \right) \right. \\
 & \qquad \left. - \left(\alpha_j^c \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds + q_{j1} \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_1^p(s) ds + \dots + q_{jr} \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_r^p(s) ds \right) \right| \right]. \tag{76}
 \end{aligned}$$

By rearranging the terms and using the triangle inequality one obtains

$$\begin{aligned}
 & \mathbb{E} \left[\left| \left(\lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^p(s) \right) \right) \right| \right] \leq \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_1^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) \right. \right. \\
 & \qquad \qquad \qquad \left. \left. - \alpha_j^c \int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds \right| \right] \\
 & \qquad \qquad \qquad + \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,1}^p(s) \right) - q_{j1} \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_1^p(s) ds \right. \right. \\
 & \qquad \qquad \qquad \vdots \\
 & \qquad \qquad \qquad \left. \left. + o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,j}^p(s) \right) - q_{jj} \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_j^p(s) ds \right. \right. \\
 & \qquad \qquad \qquad \vdots \\
 & \qquad \qquad \qquad \left. \left. + o(n) \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,r}^p(s) \right) \right. \right. \\
 & \qquad \qquad \qquad \left. \left. - q_{jr} \int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_r^p(s) ds \right| \right]. \tag{77}
 \end{aligned}$$

Note that $\int_{\mathcal{Z}} \gamma_{z,z'}^{j,p}(x) \mu_i^p(s) ds = \mathbb{E} \left[\gamma_{z,z'}^{j,p} \left(Y_{m,i}^p(s) \right) \right]$ for $m \in C_i^p$ and $\int_{\mathcal{Z}} \gamma_{z,z'}^{j,c}(x) \mu_j^c(s) ds = \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{n,j}^c(s) \right) \right]$ for $n \in C_j^c$. Then, by using the exchangeability of $Y_{n,j}^c(t)$ for $n \in C_j^c$, one finds

$$\begin{aligned} \mathbb{E} \left[\left| \left(\lambda_{j,z,z'}^p \left(v_{n,j}^{p,N}(s) \right) - \lambda_{j,z,z'}^p \left(v_j^p(s) \right) \right) \right| \right] &\leq \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) - \alpha_j^c \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right] \right| \right] \\ &+ \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,1}^p(s) \right) - q_{j1} \mathbb{E} \left[\gamma_{z,z'}^{j,p} \left(Y_{m,1}^p(s) \right) \right] \right| \right] \\ &\vdots \\ &+ o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,j}^p(s) \right) - q_{jj} \mathbb{E} \left[\gamma_{z,z'}^{j,p} \left(Y_{m,j}^p(s) \right) \right] \\ &\vdots \\ &+ o(n) \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p} \left(X_{m,r}^p(s) \right) - q_{jr} \mathbb{E} \left[\gamma_{z,z'}^{j,p} \left(Y_{m,r}^p(s) \right) \right] \Bigg]. \end{aligned} \tag{78}$$

Observe that there are $r + 1$ terms on the right-hand side of the last inequality. Let us start with the first expectation. By adding and subtracting terms one gets

$$\begin{aligned} &\mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) - \alpha_j^c \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right] \right| \right] \\ &\leq \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \left(\gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) - \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right) \right| \right] \\ &\quad + \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) - \alpha_j^c \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right] \right| \right]. \end{aligned} \tag{79}$$

Note that, by the Lipschitz property of the functions $\gamma_{z,z'}^{j,c}$, one finds that

$$\mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \left(\gamma_{z,z'}^{j,c} \left(X_{m,j}^c(s) \right) - \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right) \right| \right] \leq \zeta_{n,j}^c L_\gamma \max_{m \in C_j^c} \mathbb{E} \| X_{m,j}^c - Y_{m,j}^c \|_s. \tag{80}$$

Moreover, using (60) together with the exchangeability of $\{ Y_{n,j}^c(s), n \in C_j^c \}$ leads to

$$\begin{aligned} &\mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) - \alpha_j^c \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right] \right| \right] \\ &\leq \mathbb{E} \left[\left| \frac{\alpha_j^c}{M_j^c(n)} \sum_{m \in \mathfrak{N}_j^c(n)} \left(\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) - \mathbb{E} \left[\gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right] \right) \right| \right] \\ &+ \mathbb{E} \left[\left| \left(o(n) - \frac{\alpha_j^c}{M_j^c(n)} \right) \sum_{m \in \mathfrak{N}_j^c(n)} \gamma_{z,z'}^{j,c} \left(Y_{m,j}^c(s) \right) \right| \right] \\ &\leq \frac{\alpha_j^c \kappa_3}{\sqrt{M_j^c(n)}} + |\zeta_{n,j}^c - \alpha_j^c| \bar{\gamma}. \end{aligned} \tag{81}$$

Now let us examine the remaining r terms on the right-hand side of (78). For simplicity of notation, denote by \mathcal{I} the left-hand side of (78). Then, by adding and subtracting terms, one gets

$$\begin{aligned}
 \mathcal{I} \leq & \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_1^p(n)} \left(\gamma_{z,z'}^{j,p}(X_{m,1}^p(s)) - \gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) \right) + \dots \right. \right. \\
 & \dots + o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \left(\gamma_{z,z'}^{j,p}(X_{m,j}^p(s)) - \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right) + \dots \\
 & \left. \left. \dots + o(n) \sum_{m \in \mathfrak{N}_r^p(n)} \left(\gamma_{z,z'}^{j,p}(X_{m,r}^p(s)) - \gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \right) \right| \right] \\
 & + \mathbb{E} \left[\left| o(n) \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) - q_{j1} \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) \right] + \dots \right. \right. \\
 & \dots + o(n) \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) - q_{jj} \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right] + \dots \\
 & \left. \left. \dots + o(n) \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) - q_{jr} \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \right] \right| \right].
 \end{aligned} \tag{82}$$

Let \mathcal{I}_1 and \mathcal{I}_2 denote respectively the first and the second expectation in the right-hand side of the inequality (82). Then, by using the Lipschitz property of the functions $\gamma_{z,z'}^{j,p}$, and recalling that, for $1 \leq i \leq r$, $M_i^p(n)$ represents the number of peripheral nodes of the i th block connecting to node n , the first expectation \mathcal{I}_1 is straightforwardly bounded as follows:

$$\begin{aligned}
 \mathcal{I}_1 \leq & L_\gamma \left(\varsigma_{n,j,1}^p \max_{m \in C_1^p} \mathbb{E} \|X_{m,1}^p - Y_{m,1}^p\|_s + \dots + \varsigma_{n,j,j}^p \max_{m \in C_j^p} \mathbb{E} \|X_{m,j}^p - Y_{m,j}^p\|_s + \dots \right. \\
 & \left. \dots + \varsigma_{n,j,r}^p \max_{m \in C_r^p} \mathbb{E} \|X_{m,r}^p - Y_{m,r}^p\|_s \right).
 \end{aligned} \tag{83}$$

Moreover, by adding and subtracting terms in \mathcal{I}_2 one obtains

$$\mathcal{I}_2 \leq \mathcal{I}_3 + \mathcal{I}_4, \tag{84}$$

where

$$\begin{aligned}
 \mathcal{I}_3 := & \mathbb{E} \left[\left| \frac{q_{j1}}{M_1^p(n)} \sum_{m \in \mathfrak{N}_1^p(n)} \left(\gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) - \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) \right] \right) \right. \right. \\
 & \vdots
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{q_{jj}}{M_j^p(n) + 1} \sum_{m \in C_j^p} \left(\gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) - \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \right] \right) \\
 & \vdots \\
 & + \frac{q_{jr}}{M_r^p(n)} \sum_{m \in \mathfrak{N}_r^p(n)} \left(\gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) - \mathbb{E} \left[\gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \right] \right) \Bigg],
 \end{aligned} \tag{85}$$

and

$$\begin{aligned}
 \mathcal{I}_4 := \mathbb{E} \Bigg[& \left(o(n) - \frac{q_{j1}}{M_1^p(n)} \right) \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) + \dots + \left(o(n) - \frac{q_{jj}}{M_j^p(n) + 1} \right) \sum_{m \in \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) + \\
 & \dots + \left(o(n) - \frac{q_{jr}}{M_r^p(n)} \right) \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \Bigg].
 \end{aligned} \tag{86}$$

First, from the triangle inequality one gets

$$\begin{aligned}
 \mathcal{I}_4 \leq \mathbb{E} \Bigg[& \left(o(n) - \frac{q_{j1}}{M_1^p(n)} \right) \sum_{m \in \mathfrak{N}_1^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,1}^p(s)) \Bigg] + \dots \\
 & \dots + \mathbb{E} \Bigg[\left(o(n) - \frac{q_{jj}}{M_j^p(n) + 1} \right) \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \Bigg] + \dots \\
 & \dots + \mathbb{E} \Bigg[\left(o(n) - \frac{q_{jr}}{M_r^p(n)} \right) \sum_{m \in \mathfrak{N}_r^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,r}^p(s)) \Bigg].
 \end{aligned} \tag{87}$$

Using (20), the exchangeability of the variables $\{Y_{n,j}^p(s), n \in C_j^p\}$, and the boundedness of the functions $\gamma_{z,z'}^{j,p}$, we easily show that the right-hand side of (87) vanishes as $N \rightarrow \infty$. Indeed, the j th term satisfies

$$\begin{aligned}
 & \left| o(n) - \frac{q_{jj}}{M_j^p(n) + 1} \right| \mathbb{E} \left| \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right| \\
 & = \left| \varsigma_{n,j,j}^p - q_{jj} \right| \frac{1}{M_j^p(n) + 1} \mathbb{E} \left| \sum_{m \in \{n\} \cup \mathfrak{N}_j^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,j}^p(s)) \right| \\
 & \leq \left| \varsigma_{n,j,j}^p - q_{jj} \right| \bar{\gamma},
 \end{aligned} \tag{88}$$

and thus goes to zero by (20). Using the same steps, one obtains for all $1 \leq i \leq r$ with $i \neq j$ that

$$\left| o(n) - \frac{q_{ji}}{M_i^p(n)} \right| \mathbb{E} \left| \sum_{m \in \mathfrak{N}_i^p(n)} \gamma_{z,z'}^{j,p}(Y_{m,i}^p(s)) \right| \leq \left| \frac{M_i^p(n)}{\text{deg}(n) + 1} - q_{ji} \right| \bar{\gamma}, \tag{89}$$

which also vanishes as $N \rightarrow \infty$ by (20); thus, so does \mathcal{I}_4 . In order to bound \mathcal{I}_3 , we use again the moment inequality (60), which straightforwardly gives us

$$I_3 \leq \frac{\theta_1 q_{j1}}{\sqrt{M_1^p(n)}} + \dots + \frac{\theta_j q_{jj}}{\sqrt{M_j^p(n) + 1}} + \dots + \frac{\theta_r q_{jr}}{\sqrt{M_r^p(n)}}, \tag{90}$$

where $\theta_1, \dots, \theta_r$ are positive constants. Now, by (75), (80), (81), (83), (88), (89), and (90), one obtains

$$\begin{aligned} \mathbb{E} \left[\left\| X_{n,j}^p - Y_{n,j}^p \right\|_t \right] &\leq K |\mathcal{E}| \int_0^t \left[\bar{\gamma} \mathbb{E} \left\| X_{n,j}^p - Y_{n,j}^p \right\|_s + \zeta_{n,j}^c L_\gamma \max_{m \in C_j^c} \mathbb{E} \left\| X_{m,j}^c - Y_{m,j}^c \right\|_s + \frac{\alpha_j^c \kappa_3}{\sqrt{M_j^c(n)}} + \right. \\ &\quad + \left| \zeta_{n,j}^c - \alpha_j^c \right| \bar{\gamma} + L_\gamma \left(\zeta_{n,j,1}^p \max_{m \in C_1^p} \mathbb{E} \left\| X_{m,1}^p - Y_{m,1}^p \right\|_s + \dots \right. \\ &\quad \left. \left. + \zeta_{n,j,j}^p \max_{m \in C_j^p} \mathbb{E} \left\| X_{m,j}^p - Y_{m,j}^p \right\|_s + \dots + \zeta_{n,j,r}^p \max_{m \in C_r^p} \mathbb{E} \left\| X_{m,r}^p - Y_{m,r}^p \right\|_s \right) \\ &\quad + \left| \zeta_{n,j,1}^p - q_{j1} \right| \bar{\gamma} + \dots + \left| \zeta_{n,j,j}^p - q_{jj} \right| \bar{\gamma} + \dots + \left| \zeta_{n,j,r}^p - q_{jr} \right| \bar{\gamma} \\ &\quad \left. + \frac{\theta_1 q_{j1}}{\sqrt{M_1^p(n)}} + \dots + \frac{\theta_j q_{jj}}{\sqrt{M_j^p(n) + 1}} + \dots + \frac{\theta_r q_{jr}}{\sqrt{M_r^p(n)}} \right] ds. \end{aligned} \tag{91}$$

Recall that $deg(n) = M_j^c(n) + \sum_{k=1}^r M_k^p(n)$ for any $n \in C_j^p$. Using this and taking the maximum over $n \in C_j^p$ and over $1 \leq j \leq r$ in (91), one gets

$$\begin{aligned} &\max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left[\left\| X_{n,j}^p - Y_{n,j}^p \right\|_t \right] \\ &\leq K |\mathcal{E}| \int_0^t \left[L_\gamma \max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left\| X_{n,j}^c - Y_{n,j}^c \right\|_s + (\bar{\gamma} + L_\gamma) \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left\| X_{n,j}^p - Y_{n,j}^p \right\|_s \right. \\ &\quad + \sum_{1 \leq j \leq r} \max_{n \in C_j^p} \left(\frac{\alpha_j^c \kappa_3}{\sqrt{M_j^c(n)}} \right) + \sum_{1 \leq j \leq r} \max_{n \in C_j^p} \left| \zeta_{n,j}^c - \alpha_j^c \right| \bar{\gamma} \\ &\quad + \sum_{1 \leq j \leq r} \left(\max_{n \in C_j^p} \left| \zeta_{n,j,1}^p - q_{j1} \right| \bar{\gamma} + \dots + \max_{n \in C_j^p} \left| \zeta_{n,j,j}^p - q_{jj} \right| \bar{\gamma} + \dots + \max_{n \in C_j^p} \left| \zeta_{n,j,r}^p - q_{jr} \right| \bar{\gamma} \right. \\ &\quad \left. + \max_{n \in C_j^p} \frac{\theta_1 q_{j1}}{\sqrt{M_1^p(n)}} + \dots + \max_{n \in C_j^p} \frac{\theta_j q_{jj}}{\sqrt{M_j^p(n) + 1}} + \dots + \max_{n \in C_j^p} \frac{\theta_r q_{jr}}{\sqrt{M_r^p(n)}} \right) \Big] ds. \end{aligned} \tag{92}$$

Adding side by side the two inequalities in (73) and (92) leads to

$$\begin{aligned} &\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left\| X_{n,j}^c - Y_{n,j}^c(t) \right\|_t + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left\| X_{n,j}^p - Y_{n,j}^p \right\|_t \\ &\leq K |\mathcal{E}| \int_0^t \left(C_1 \max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left\| X_{n,j}^c - Y_{n,j}^c(t) \right\|_s + C_2 \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left\| X_{n,j}^p - Y_{n,j}^p \right\|_s + C_3(N) \right) ds, \end{aligned} \tag{93}$$

where, with a slight abuse of notation, the constants C_1 , C_2 and the function $C_3(N)$ are defined by

$$C_1 := \bar{\gamma} + 2L_\gamma, \\ C_2 := \bar{\gamma} + 2L_\gamma,$$

and

$$C_3(N) := \sum_{1 \leq j \leq r} \left(\bar{\gamma} \max_{n \in C_j^c} |\varrho_{n,j}^c - p_j^c| + \bar{\gamma} \max_{n \in C_j^c} |\varrho_{n,j}^c - p_j^p| + \max_{n \in C_j^c} \frac{\kappa_1 p_j^c}{\sqrt{M_j^c(n) + 1}} + \max_{n \in C_j^c} \frac{\kappa_2 p_j^p}{\sqrt{M_j^p(n)}} \right) \\ + \sum_{1 \leq j \leq r} \max_{n \in C_j^p} \left(\frac{\alpha_j^c \kappa_3}{\sqrt{M_j^c(n)}} \right) + \sum_{1 \leq j \leq r} \max_{n \in C_j^p} |\varsigma_{n,j}^c - \alpha_j^c| \bar{\gamma} \\ + \sum_{1 \leq j \leq r} \left(\max_{n \in C_j^p} |\varsigma_{n,j,1}^p - q_{j1}| \bar{\gamma} + \dots + \max_{n \in C_j^p} |\varsigma_{n,j,j}^p - q_{jj}| \bar{\gamma} + \dots + \max_{n \in C_j^p} |\varsigma_{n,j,r}^p - q_{jr}| \bar{\gamma} \right. \\ \left. + \max_{n \in C_j^p} \frac{\theta_1 q_{j1}}{\sqrt{M_1^p(n)}} + \dots + \max_{n \in C_j^p} \frac{\theta_j q_{jj}}{\sqrt{M_j^p(n) + 1}} + \dots + \max_{n \in C_j^p} \frac{\theta_r q_{jr}}{\sqrt{M_r^p(n)}} \right).$$

Therefore, applying Grönwall’s lemma to (93) gives

$$\max_{1 \leq j \leq r} \max_{n \in C_j^c} \mathbb{E} \left[\left\| X_{n,j}^c - Y_{n,j}^c \right\|_t \right] + \max_{1 \leq j \leq r} \max_{n \in C_j^p} \mathbb{E} \left[\left\| X_{n,j}^p - Y_{n,j}^p \right\|_t \right] \leq K |\mathcal{E}| C_3(N) t \exp \left\{ \int_0^t C_4 ds \right\}, \tag{94}$$

with $C_4 = C_1 + C_2$. Finally, Condition 4.1 ensures that $C_3(N) \rightarrow 0$ as $N \rightarrow \infty$, which proves (61).

We are now ready to conclude the proof. First, Theorem 4.1 ensures the uniqueness of the solution of the limiting stochastic differential equation (14). In addition, the relation in (25) shows that the solution is continuous with respect to the initial condition. Therefore, the process $Y(t) = ((Y_n^c(t), Y_m^p(t)), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$ is $\mu_1^c \otimes \mu_1^p \otimes \dots \otimes \mu_r^c \otimes \mu_r^p$ -multi-chaotic since the initial condition $Y(0) = X(0)$ is multi-exchangeable and $v^{1,c} \otimes v^{1,p} \otimes \dots \otimes v^{r,c} \otimes v^{r,p}$ -multi-chaotic. Then, by the relation in (61), we conclude that the convergence in (59) holds and the sequence of processes $((X_n^{c,N}(t), X_m^{p,N}(t), t \geq 0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$ is also $P_{\bar{X}} = \mu_1^c \otimes \mu_1^p \otimes \dots \otimes \mu_r^c \otimes \mu_r^p$ -multi-chaotic, which concludes the proof. \square

5.2. Laws of large numbers

The following laws of large numbers are immediate consequences of Theorem 5.1.

Corollary 5.1. *Suppose that the conditions of Theorem 5.1 hold. Define $\mu_j^c := \mathcal{L}(\bar{X}_{n,j}^c)$, $\mu_j^p := \mathcal{L}(\bar{X}_{m,j}^p)$ for $1 \leq j \leq r$, where $((\bar{X}_{n,j}^c(t), \bar{X}_{m,j}^p(t), t \geq 0), 1 \leq j \leq r)$ is the solution of the McKean–Vlasov limiting system in (14) with initial distribution $v^{1,c} \otimes v^{1,p} \dots \otimes v^{r,c} \otimes v^{r,p}$. Then, for each $1 \leq j \leq r$, as $N \rightarrow \infty$,*

$$\mu_j^{c,N} = \frac{1}{N_j^c} \sum_{n \in C_j^c} \delta_{X_{n,j}^c} \rightarrow \mu_j^c \tag{95}$$

and

$$\mu_j^{p,N} = \frac{1}{N_j^p} \sum_{n \in C_j^p} \delta_{X_{n,j}^p} \rightarrow \mu_j^p, \tag{96}$$

for the weak topology on $\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}))$ with $\mathcal{D}([0, T], \mathcal{Z})$ endowed with the Skorokhod topology.

Proof. We prove (96); the proof of (95) is similar. Let

$$\bar{\mu}_j^{N,p} := \frac{1}{N_j^p} \sum_{n \in C_j^p} \delta_{\bar{X}_{n,j}^p},$$

and recall that the bounded-Lipschitz metric d_{BL} metrizes the weak convergence on $\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}))$. Therefore, to prove the convergence in (96), it suffices to prove that $d_{BL}(\mu_j^{p,N}, \bar{\mu}_j^{p,N}) \Rightarrow 0$ and that $\bar{\mu}_j^{N,p} \Rightarrow \mu_j^p$. First note that for all $1 \leq j \leq r$,

$$\begin{aligned} \mathbb{E} \left[d_{BL}(\mu_j^{p,N}, \bar{\mu}_j^{p,N}) \right] &= \mathbb{E} \left[\sup_{g \in Lip(\mathcal{Z})} \left| \langle \mu_j^{p,N}, g \rangle - \langle \bar{\mu}_j^{p,N}, g \rangle \right| \right] \\ &= \mathbb{E} \left[\sup_{g \in Lip(\mathcal{Z})} \left| \frac{1}{N_j^p} \sum_{n \in C_j^p} (g(X_{n,j}^p) - g(\bar{X}_{n,j}^p)) \right| \right] \\ &\leq \frac{1}{N_j^p} \sum_{n \in C_j^p} \mathbb{E} \left[|X_{n,j}^p - \bar{X}_{n,j}^p|_T \right] \\ &\leq \max_{n \in C_j^p} \mathbb{E} |X_{n,j}^p - \bar{X}_{n,j}^p|_T, \end{aligned} \tag{97}$$

which goes to zero according to (59). Thus, $d_{BL}(\mu_j^{p,N}, \bar{\mu}_j^{p,N}) \Rightarrow 0$ as $N \rightarrow \infty$. It remains to show that $\bar{\mu}_j^{N,p} \Rightarrow \mu_j^p$ as $N \rightarrow \infty$. Since the stochastic processes $\{\bar{X}_{n,j}^p, n \in C_j^p\}$ are i.i.d., for any continuous and bounded function $g \in C_b(\mathcal{Z})$ one finds that

$$\begin{aligned} \mathbb{E} \left(\langle \bar{\mu}_j^{p,N}, g \rangle - \langle \mu_j^p, g \rangle \right)^2 &= \mathbb{E} \left(\frac{1}{N_j^p} \sum_{n \in C_j^p} (g(\bar{X}_{n,j}^p) - \langle \mu_j^p, g \rangle) \right)^2 \\ &= \mathbb{E} \left(\frac{1}{(N_j^p)^2} \sum_{n \in C_j^p} (g(\bar{X}_{n,j}^p) - \langle \mu_j^p, g \rangle)^2 \right) \\ &\leq \frac{1}{N_j^p} 4 \|g\|_\infty^2, \end{aligned} \tag{98}$$

which goes to zero given the boundedness of g . Therefore, $\bar{\mu}_j^{p,N}$ converges weakly to μ_j^p as $N \rightarrow \infty$. Thus, combining the two convergence results, we conclude that $\mu_j^{p,N}$ converges weakly to μ_j^p as $N \rightarrow \infty$. The corollary is proved. \square

6. Large deviations

We investigate here the large deviations principles of the interacting particle system introduced in Section 2 over finite time durations. For the sake of simplicity, we restrict ourselves to the case of block-structured graphs where the blocks are cliques, i.e., complete subgraphs, and the peripheral subgraph is complete, that is, all peripheral nodes in the system are connected; see, e.g., Figure 1. The first main result of this section is Theorem 6.1, which states the large deviations principle of the vector of empirical measures. The second main result is Theorem 6.2, which states the large deviations principle of the corresponding vector of empirical processes. The approach we take to establish these results is based on a generalization to the multi-class setting of the classical approach developed in [24] and adapted in [49] to the context of jump processes. One might also consult [32, 33] for an alternative approach.

Let us first introduce the assumptions under which the results of this section hold.

Assumption 6.1.

1. The peripheral subgraph is complete; that is, for any two peripheral nodes $n, m \in \bigcup_{1 \leq j \leq r} C_j^p$, there exists an edge $(n, m) \in \Xi$ connecting n and m .
2. The r blocks of the graphs are cliques; that is, for any two nodes $n, m \in C_j^p$ of the same block $1 \leq j \leq r$, there exists an edge $(n, m) \in \Xi$ connecting n and m .
3. The mappings $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ introduced in (16) and (17) are uniformly bounded away from zero; that is, there exists $c > 0$ such that, for all $v, \mu_1, \dots, \mu_r \in \mathcal{M}_1(\mathcal{Z})$ and all $(z, z') \in \mathcal{E}$, we have $\lambda_{j,z,z'}^c(v, \mu_j) \geq c$ and $\lambda_{j,z,z'}^p(v, \mu_1, \dots, \mu_r) \geq c$.
4. As $N \rightarrow \infty$,

$$\frac{N_j}{N} \rightarrow \alpha_j \tag{99}$$

for some $\alpha_j \in (0, 1)$, where we recall that N_j is the number of nodes in the j th block and N_j^c (resp. N_j^p) is the number of central (resp. peripheral) nodes in the j th block.

Remark 6.1.

1. From (16) and (17), together with Remark 4.1, the functions $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ are Lipschitz.
2. Since $\mathcal{M}_1(\mathcal{Z})$ is compact and the rate functions $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ are continuous and Lipschitz, the rates are uniformly bounded from above; that is, there exists a constant $C < \infty$ such that for all $v, \mu_1, \dots, \mu_r \in \mathcal{M}_1(\mathcal{Z})$, and all $(z, z') \in \mathcal{E}$, we have $\lambda_{j,z,z'}^c(v, \mu_j) \leq C$ and $\lambda_{j,z,z'}^p(v, \mu_1, \dots, \mu_r) \leq C$.
3. For ease of reading, we have omitted subscripts indicating the dependence of the rate functions $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ on the proportions $a_1, b_1, a, b_1, \dots, b_r$.
4. We use again throughout this section the convention that N goes to infinity when both $\min_{1 \leq j \leq r} N_j^c$ and $\min_{1 \leq j \leq r} N_j^p$ go to infinity.

5. We emphasize that for simplicity, the results obtained in this section hold under Assumption 6.1, which describes a special case of the class of models given by Condition 4.1; that is, we suppose here that each block is a clique and the peripheral subgraph is complete.

Let $\mathbb{M}^N \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ denote the vector of empirical measures defined by

$$\begin{aligned} \mathbb{M}^N &:= \left(\mathbb{M}_1^{c,N}, \mathbb{M}_1^{p,N}, \dots, \mathbb{M}_r^{c,N}, \mathbb{M}_r^{p,N} \right) \\ &= \left(\frac{1}{N_1^c} \sum_{n \in C_1^c} \delta_{X_{n,1}^c}, \frac{1}{N_1^p} \sum_{n \in C_1^p} \delta_{X_{n,1}^p}, \dots, \frac{1}{N_r^c} \sum_{n \in C_r^c} \delta_{X_{n,r}^c}, \frac{1}{N_r^p} \sum_{n \in C_r^p} \delta_{X_{n,r}^p} \right), \end{aligned} \tag{100}$$

where $X^N = (X_{n,j}^c, X_{m,j}^p; n \in C_j^c, m \in C_j^p; 1 \leq j \leq r) \in \mathcal{D}([0, T], \mathcal{Z}^N)$ denotes the full description of the N particles and $\mathbb{M}_j^{c,N}$ (resp. $\mathbb{M}_j^{p,N}$) is the empirical measure of the central (resp. peripheral) nodes of the j th block, for $1 \leq j \leq r$. With a slight abuse of notation, denote by $G_N : \mathcal{D}([0, T], \mathcal{Z}^N) \rightarrow (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ the mapping that takes the full description X^N to the empirical measures vector \mathbb{M}^N , that is,

$$\begin{aligned} G_N : (X_{n,j}^c, X_{m,j}^p; n \in C_j^c, m \in C_j^p; 1 \leq j \leq r) &\rightarrow \\ &\left(\frac{1}{N_1^c} \sum_{n \in C_1^c} \delta_{X_{n,1}^c}, \frac{1}{N_1^p} \sum_{n \in C_1^p} \delta_{X_{n,1}^p}, \dots, \frac{1}{N_r^c} \sum_{n \in C_r^c} \delta_{X_{n,r}^c}, \frac{1}{N_r^p} \sum_{n \in C_r^p} \delta_{X_{n,r}^p} \right). \end{aligned}$$

Thus, $\mathbb{M}^N = G_N(X^N)$. Denote by $\mathbb{P}_{z^N}^N$ the law of X^N with initial condition $z^N = (z_{n,j}^c, z_{m,j}^p; n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$. Note that the distribution of the empirical vector \mathbb{M}^N depends on the initial condition only through its empirical vector, defined by

$$\begin{aligned} \nu_N &:= \left(\nu_N^{1,c}, \nu_N^{1,p}, \dots, \nu_N^{r,c}, \nu_N^{r,p} \right) = \\ &\left(\frac{1}{N_1^c} \sum_{n \in C_1^c} \delta_{z_{n,1}^c}, \frac{1}{N_1^p} \sum_{n \in C_1^p} \delta_{z_{n,1}^p}, \dots, \frac{1}{N_r^c} \sum_{n \in C_r^c} \delta_{z_{n,r}^c}, \frac{1}{N_r^p} \sum_{n \in C_r^p} \delta_{z_{n,r}^p} \right). \end{aligned} \tag{101}$$

Moreover, denote by $P_{\nu_N}^N$ the distribution of \mathbb{M}^N , which is the pushforward of $\mathbb{P}_{z^N}^N$ under the mapping G_N ; that is, $P_{\nu_N}^N = \mathbb{P}_{z^N}^N \circ G_N^{-1}$.

Let us now introduce the $(\mathcal{M}_1(\mathcal{Z}))^{2r}$ -valued vector of empirical processes

$$\begin{aligned} \mu^N : t \in [0, T] &\longrightarrow \mu^N(t) = \left(\mu_1^{c,N}(t), \mu_1^{p,N}(t), \dots, \mu_r^{c,N}(t), \mu_r^{p,N}(t) \right) \\ &= \left(\frac{1}{N_1^c} \sum_{n \in C_1^c} \delta_{X_{n,1}^c(t)}, \frac{1}{N_1^p} \sum_{n \in C_1^p} \delta_{X_{n,1}^p(t)}, \dots, \frac{1}{N_r^c} \sum_{n \in C_r^c} \delta_{X_{n,r}^c(t)}, \frac{1}{N_r^p} \sum_{n \in C_r^p} \delta_{X_{n,r}^p(t)} \right), \end{aligned} \tag{102}$$

and denote by γ_N the corresponding mapping that takes a full description $X^N \in \mathcal{D}([0, T], \mathcal{Z}^N)$ of the N particles of the system to the empirical process vector μ^N , that is,

$$\gamma_N : (X_{n,j}^c, X_{m,j}^p; n \in C_j^c, m \in C_j^p; 1 \leq j \leq r) \in \mathcal{D}([0, T], \mathcal{Z}^N) \longrightarrow \mu^N : [0, T] \longrightarrow (\mathcal{M}_1(\mathcal{Z}))^{2r}.$$

Observe that $\mu^N(0) = \nu_N$ and that $\mu^N(t)$ is the projection $\pi_t(\mathbb{M}^N)$ at time t , that is,

$$\mu^N = \pi(\mathbb{M}^N) = \pi(G_N(X^N)) = \gamma_N(X^N),$$

where the notation π denotes, again with a slight abuse of notation, both the vector projection

$$\pi : (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r} \rightarrow (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$$

and the component projection

$$\pi : \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})) \rightarrow \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})).$$

Finally, denote by $p_{\nu_N}^N$ the distribution of μ^N , which is the pushforward $p_{\nu_N}^N = \mathbb{P}_{z_N}^N \circ \gamma_N^{-1}$. Note that, since $\mu^N = \pi(M^N)$, we can also write $p_{\nu_N}^N$ as the pushforward $p_{\nu_N}^N = P_{\nu_N}^N \circ \pi^{-1}$.

The goal of this section is to study the large deviations principles for the sequences of probability measures $(P_{\nu_N}^N, N \geq 1)$ and $(p_{\nu_N}^N, N \geq 1)$. The two main results are Theorem 6.1 and Theorem 6.2.

6.1. Large deviations principle for the empirical measure vector

We start by investigating the large deviations principle of the sequence $(P_{\nu_N}^N, N \geq 1)$. To this end, we first establish the result in the non-interacting case. Then, through the Radon–Nikodym derivative, one uses the Laplace–Varadhan principle to deduce the case with interactions.

Let us first describe the hypothetical non-interacting system. Suppose that all the nodes are independent of each other and that the color of each node changes with a constant rate equal to 1 for all allowed transitions $(z, z') \in \mathcal{E}$, while all other transition rates are zero. Denote by P_{z_0} the marginal law on $\mathcal{D}([0, T], \mathcal{Z})$ of this process with initial condition z_0 . Thus, P_{z_0} is the unique solution to the martingale problem in $\mathcal{D}([0, T], \mathcal{Z})$ associated with the generator \mathcal{L}^0 operating on bounded measurable functions ϕ on \mathcal{Z} according to

$$\mathcal{L}^0 \phi(z) := \sum_{z' : (z, z') \in \mathcal{E}} 1 \cdot (\phi(z') - \phi(z))$$

and the initial condition z_0 . Given that the transition rates are upper-bounded and that

$$\sup_{z \in \mathcal{Z}} \sum_{z' : (z, z') \in \mathcal{E}} |z' - z| < F(1 + z)$$

for some constant F , there exists a unique solution to the martingale problem for (\mathcal{L}^0, z_0) (cf. [31, Problem 4.11.15]).

For any $\eta, \rho_1, \dots, \rho_r$ in $\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))$, let $R_{z_0}^c(\eta, \rho_j)$ be the unique solution to the martingale problem in $\mathcal{D}([0, T], \mathcal{Z})$ associated with the time-varying generator

$$\mathcal{L}_{\eta(t), \rho_j(t)}^c \phi(z) := \sum_{z' : (z, z') \in \mathcal{E}} \lambda_{j, z, z'}^c(\eta(t), \rho_j(t)) (\phi(z') - \phi(z)) \tag{103}$$

and the initial condition z_0 . Similarly, let $R_{z_0}^p(\eta, \rho_1, \dots, \rho_r)$ be the unique solution to the martingale problem in $\mathcal{D}([0, T], \mathcal{Z})$ associated with the time-varying generator

$$\mathcal{L}_{\eta(t), \rho_1(t), \dots, \rho_r(t)}^p \phi(z) := \sum_{z' : (z, z') \in \mathcal{E}} \lambda_{j, z, z'}^p(\eta(t), \rho_1(t), \dots, \rho_r(t)) (\phi(z') - \phi(z)) \tag{104}$$

and the initial condition z_0 . Again by the upper-boundedness of $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$, the uniqueness of $R_{z_0}^c(\eta, \rho_j)$ and $R_{z_0}^p(\eta, \rho_1, \dots, \rho_r)$ follows (see again [31, Problem 4.11.15]). Therefore, the density of $R_{z_0}^c(\eta, \rho_j)$ and $R_{z_0}^p(\eta, \rho_1, \dots, \rho_r)$ with respect to P_{z_0} can be written as follows (see [49, Equation (2.4)]):

$$\frac{dR_{z_0}^c(\eta, \rho_j)}{dP_{z_0}}(x) = \exp\{h_1(x, \eta, \rho_j)\} \quad \text{and} \quad \frac{dR_{z_0}^p(\eta, \rho_1, \dots, \rho_r)}{dP_{z_0}}(x) = \exp\{h_2(x, \eta, \rho_1, \dots, \rho_r)\}, \tag{105}$$

where

$$h_1(x, \eta, \rho_j) := \sum_{0 \leq t \leq T} \mathbb{1}_{\{x_t \neq x_{t-}\}} \log \left(\lambda_{j,x_{t-},x_t}^c(\eta(t-), \rho_j(t-)) \right) - \int_0^T \left(\sum_{z: (x_t,z) \in \mathcal{E}} \lambda_{j,x_t,z}^c(\eta(t), \rho_j(t)) - 1 \right) dt \tag{106}$$

and

$$h_2(x, \eta, \rho_1, \dots, \rho_r) := \sum_{0 \leq t \leq T} \mathbb{1}_{\{x_t \neq x_{t-}\}} \log \left(\lambda_{j,x_{t-},x_t}^p(\eta(t-), \rho_1(t-), \dots, \rho_r(t-)) \right) - \int_0^T \left(\sum_{z: (x_t,z) \in \mathcal{E}} \lambda_{j,x_t,z}^p(\eta(t), \rho_1(t), \dots, \rho_r(t)) - 1 \right) dt. \tag{107}$$

Consider now a system of N non-interacting particles where the law of the n th particle is P_{z_n} with initial condition z_n . The law of such a system is the product distribution $\mathbb{P}_{z^N}^{0,N} = \otimes_{n=1}^N P_{z_n}$. Moreover, the distribution of the corresponding empirical vector is given by $P_{\nu^N}^{0,N} = \mathbb{P}_{z^N}^{0,N} \circ G_N^{-1}$ where ν_N is the initial empirical vector (101). Therefore, by applying an analogue of the Cameron–Martin–Girsanov formula for stochastic integrals with respect to point processes (see e.g. [27, Lemma 3.7] or [49, Equation (2.8)]), one can compute the Radon–Nikodym derivative $dP_{\nu^N}^N/dP_{\nu^N}^{0,N}$ at any $\mathbf{Q} = (Q_1^c, Q_1^p, \dots, Q_r^c, Q_r^p) \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ as follows:

$$\begin{aligned} \frac{dP_{\nu^N}^N}{dP_{\nu^N}^{0,N}}(\mathbf{Q}) &= \exp \left\{ \sum_{j=1}^r \left[N_j^c \int_{D([0,T], \mathcal{Z})} h_1(x, \pi(Q_j^c), \pi(Q_j^p)) Q_j^c(dx) \right. \right. \\ &\quad \left. \left. + N_j^p \int_{D([0,T], \mathcal{Z})} h_2(x, \pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p)) Q_j^p(dx) \right] \right\} \\ &= \exp \{Nh^N(\mathbf{Q})\}, \end{aligned} \tag{108}$$

with

$$\pi(\mathbf{Q}) = (\pi(Q_1^c), \pi(Q_1^p), \dots, \pi(Q_r^c), \pi(Q_r^p))$$

being the vector containing the component projections $\pi(Q_j^\iota) \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))$ for $1 \leq j \leq r$ and $\iota \in \{c, p\}$, and where

$$h^N(\mathbf{Q}) := \sum_{j=1}^r \left[\frac{N_j^c}{N} \int_{D([0, T], \mathcal{Z})} h_1(x, \pi(Q_j^c), \pi(Q_j^p)) Q_j^c(dx) + \frac{N_j^p}{N} \int_{D([0, T], \mathcal{Z})} h_2(x, \pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p)) Q_j^p(dx) \right]. \tag{109}$$

Note that under Assumption 6.1, the sequence of functions $\{h^N\}_{N \geq 1}$ converges, as $N \rightarrow \infty$, towards the function h given by

$$h(\mathbf{Q}) := \sum_{j=1}^r \left[\alpha_j p_j^c \int_{D([0, T], \mathcal{Z})} h_1(x, \pi(Q_j^c), \pi(Q_j^p)) Q_j^c(dx) + \alpha_j p_j^p \int_{D([0, T], \mathcal{Z})} h_2(x, \pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p)) Q_j^p(dx) \right]. \tag{110}$$

We now introduce the necessary spaces and topologies following the notation of [12, 49]. Consider the Polish space (\mathcal{X}, d) where

$$\mathcal{X} := \left\{ x \in \mathcal{D}([0, T], \mathcal{Z}) \mid \sum_{0 \leq t \leq T} \mathbb{1}_{x_t \neq x_{t-}} < +\infty, \right. \\ \left. \text{and for each } t \in (0, T] \text{ with } x_t \neq x_{t-}, \text{ we have } (x(t-), x(t)) \in \mathcal{E} \right\},$$

and the metric d is defined by

$$d(x, y) := d_{Sko}(x, y) + |\varphi(x) - \varphi(y)|, \quad x, y \in \mathcal{X}, \tag{111}$$

with $\varphi(x) := \sum_{0 \leq t \leq T} \mathbb{1}_{x_t \neq x_{t-}}$ denoting the number of jumps and d_{Sko} standing for the Skorokhod complete metric (see [10, Section 12]). For this topology, the function φ is continuous, and two paths are close to each other if they have the same number of jumps and if they are Skorokhod-close [49, p. 299]. For any function $f : \mathcal{X} \rightarrow \mathbb{R}$, define

$$\|f\|_\varphi := \sup_{x \in \mathcal{X}} \frac{f(x)}{1 + \varphi(x)}, \tag{112}$$

and write

$$C_\varphi(\mathcal{X}) := \{f : \mathcal{X} \rightarrow \mathbb{R} \text{ is continuous and } \|f\|_\varphi < \infty\}, \tag{113}$$

$$\mathcal{M}_{1,\varphi}(\mathcal{X}) := \left\{ Q \in \mathcal{M}_1(\mathcal{X}) \mid \int_{\mathcal{X}} \varphi dQ < +\infty \right\}. \tag{114}$$

We endow the set $\mathcal{M}_{1,\varphi}(\mathcal{X})$ with the weak-* topology $\sigma(\mathcal{M}_{1,\varphi}(\mathcal{X}), C_\varphi(\mathcal{X}))$, the weakest topology under which $Q_N \rightarrow Q$ as $N \rightarrow +\infty$ if and only if

$$\int_{\mathcal{X}} f dQ_N \rightarrow \int_{\mathcal{X}} f dQ \quad \text{for each } f \in C_\varphi(\mathcal{X}).$$

For a measure $\nu = (\nu^{1,c}, \nu^{1,p}, \dots, \nu^{r,c}, \nu^{r,p}) \in (\mathcal{M}_1(\mathcal{Z}))^{2r}$ we define, for all $1 \leq j \leq r$ and $\iota \in \{c, p\}$, the mixture

$$dP_{j,\iota}(x) := \sum_{z_0 \in \mathcal{Z}} \nu^{j,\iota}(z_0) dP_{z_0}(x).$$

Moreover, let $R^c(\eta, \rho_j)$ and $R^p(\eta, \rho_1, \dots, \rho_r)$ be the mixtures given by

$$\begin{aligned} dR^c(\eta, \rho_j)(x) &:= \sum_{z_0 \in \mathcal{Z}} \nu^{j,c}(z_0) dR_{z_0}^c(\eta, \rho_j)(x), \\ dR^p(\eta, \rho_1, \dots, \rho_r)(x) &:= \sum_{z_0 \in \mathcal{Z}} \nu^{j,p}(z_0) dR_{z_0}^p(\eta, \rho_1, \dots, \rho_r)(x). \end{aligned} \tag{115}$$

Finally, let us introduce the relative entropy $H : \mathcal{M}_{1,\varphi}(\mathcal{X}) \rightarrow [0, +\infty]$ of Q with respect to P as follows:

$$H(Q|P) := \begin{cases} \int_{\mathcal{X}} \log \left(\frac{dQ}{dP} \right) dQ & \text{if } Q \ll P, \\ +\infty & \text{otherwise.} \end{cases} \tag{116}$$

We are now ready to state the large deviations principle for the sequence $(P_{\nu_N}^N, N \geq 1)$.

Theorem 6.1. *Let the space $\mathcal{M}_{1,\varphi}(\mathcal{X})$ be equipped with the weak- $*$ topology $\sigma(\mathcal{M}_{1,\varphi}(\mathcal{X}), C_\varphi(\mathcal{X}))$. Moreover, suppose that the initial condition $\nu_N \rightarrow \nu$ weakly as $N \rightarrow \infty$. Then the sequence $(P_{\nu_N}^N, N \geq 1)$ satisfies the large deviations principle in the space $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, endowed with the product topology, with speed N and the good rate function $I(\mathbf{Q}) = L(\mathbf{Q}) - h(\mathbf{Q})$, where the function $h(\mathbf{Q})$ is given by (110) and $L : (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r} \rightarrow [0, \infty]$ is defined as*

$$L(\mathbf{Q}) := \alpha_1 p_1^c J^{1,c}(\mathcal{Q}_1^c) + \alpha_1 p_1^p J^{1,p}(\mathcal{Q}_1^p) + \dots + \alpha_r p_r^c J^{r,c}(\mathcal{Q}_r^c) + \alpha_r p_r^p J^{r,p}(\mathcal{Q}_r^p), \tag{117}$$

with, for each $1 \leq j \leq r$, $\iota \in \{c, p\}$, and $Q \in \mathcal{M}_{1,\varphi}(\mathcal{X})$,

$$J^{j,\iota}(Q) := \sup_{f \in C_\varphi(\mathcal{X})} \left[\int_{\mathcal{X}} f dQ - \sum_{z_0 \in \mathcal{Z}} \nu^{j,\iota}(z_0) \log \int_{\mathcal{X}} e^f dP_{z_0} \right], \tag{118}$$

and α_j, p_j^c, p_j^p being given in (99) and (18). Furthermore, for each $\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, the rate function $I(\mathbf{Q})$ admits the representation

$$\begin{cases} \sum_{j=1}^r \left[\alpha_j p_j^c H \left(\mathcal{Q}_j^c | R^c(\pi(\mathcal{Q}_j^c), \pi(\mathcal{Q}_j^p)) \right) \right. \\ \quad \left. + \alpha_j p_j^p H \left(\mathcal{Q}_j^p | R^p(\pi(\mathcal{Q}_j^c), \pi(\mathcal{Q}_1^p), \dots, \pi(\mathcal{Q}_r^p)) \right) \right] & \text{if } \mathbf{Q} \circ \pi_0^{-1} = \nu, \\ +\infty & \text{otherwise.} \end{cases} \tag{119}$$

Remark 6.2. This is a generalization of [49, Theorem 2.1] to our multi-population setting. Also, while [49] studied the case where $z_n = z_0$ for some fixed z_0 , so that $v_N = \delta_{z_0}$, we consider, as in [12, Theorem 3.1], more general initial conditions, provided that the initial empirical vector v_N converges weakly towards $v = (v^{1,c}, v^{1,p}, \dots, v^{r,c}, v^{r,p})$. Moreover, similarly to [12], we consider here the case where not all transitions are allowed, but only those in \mathcal{E} , the set of directed edges in the graph $(\mathcal{Z}, \mathcal{E})$.

Note that, from Definition 5.1, the weak convergence of the initial empirical vector v_N towards v amounts to the assertion that the initial conditions $(X_{n,j}^c(0), X_{m,j}^p(0), n \in C_j^c, m \in C_j^p; 1 \leq j \leq r)$ are $v^{1,c} \otimes v^{1,p} \dots v^{r,c} \otimes v^{r,p}$ -multi-chaotic (cf. [61]).

Proof of Theorem 6.1. The proof of Theorem 6.1 is based on the generalization of Sanov’s theorem for empirical measures on Polish spaces due to Dawson and Gärtner [24], the Girsanov transformation, and the Laplace–Varadhan principle [64]. We proceed through several lemmas. We follow [49, Theorem 2.1] and [12, Theorem 3.1].

Large deviations principle for the non-interacting case. We first establish a large deviations principle in the non-interacting case. □

Lemma 6.1. *Suppose that the initial condition v_N converges towards v weakly as $N \rightarrow \infty$. Let $\mathcal{M}_{1,\varphi}(\mathcal{X})$ be endowed with the weak-* topology $\sigma(\mathcal{M}_{1,\varphi}(\mathcal{X}), C_\varphi(\mathcal{X}))$. Then the sequence $(P_{v_N}^{0,N}, N \geq 1)$ satisfies a large deviations principle in $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, endowed with the product topology, with speed N and the action functional $L : (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r} \rightarrow [0, \infty]$ given by (117).*

Proof. Fix a given block $1 \leq j \leq r$. Denote by

$$\left(P_{v_N^{j,c}}^{0,N_j^c}, N_j^c \geq 1 \right) \quad \text{and} \quad \left(P_{v_N^{j,p}}^{0,N_j^p}, N_j^p \geq 1 \right)$$

the sequences of probability distributions of the local empirical measures $\mathbb{M}_j^{N,c}$ and $\mathbb{M}_j^{N,p}$ of the central and peripheral nodes, respectively, of the j th block. Note that in the non-interacting case, the transition rate from any state to any other state is bounded by 1. Therefore, the family of probability measures $\{P_z : z \in \mathcal{Z}\}$ is a subset of $\mathcal{M}_{1,\varphi}(\mathcal{X})$. Moreover, for any continuous function $F \in C_\varphi(\mathcal{X})$, the integral $\int F(y)P_{z_0}(dy)$ depends continuously upon z_0 , and then $\{P_{z_0} : z_0 \in \mathcal{Z}\}$ is a Feller continuous family of probability measures on \mathcal{X} . Now, since $v_N^{j,c} \rightarrow v^{j,c}$ and $v_N^{j,p} \rightarrow v^{j,p}$ as $N \rightarrow \infty$, by applying the generalization of Sanov’s theorem [24, Theorem 3.5], we find that both the sequences $\left(P_{v_N^{j,c}}^{0,N_j^c}, N_j^c \geq 1 \right)$ and $\left(P_{v_N^{j,p}}^{0,N_j^p}, N_j^p \geq 1 \right)$ satisfy the large deviations principle in $\mathcal{M}_{1,\varphi}(\mathcal{X})$ endowed with the weak-* topology $\sigma(\mathcal{M}_{1,\varphi}(\mathcal{X}), C_\varphi(\mathcal{X}))$, with speeds N_j^c and N_j^p , respectively, and good rate functions $J^{j,c}(Q)$ and $J^{j,p}(Q)$ defined by (118). Let $\mathcal{K}_1^c, \mathcal{K}_1^p, \dots, \mathcal{K}_r^c, \mathcal{K}_r^p \in \mathcal{B}(\mathcal{M}_{1,\varphi}(\mathcal{X}))$ be closed Borelian sets. By independence, one has

$$P_{v_N}^{0,N} \left\{ M^N \in \prod_{j=1}^r (\mathcal{K}_j^c \times \mathcal{K}_j^p) \right\} = \prod_{j=1}^r \left(P_{v_N^{j,c}}^{0,N_j^c} \left\{ M_j^{N,c} \in \mathcal{K}_j^c \right\} \times P_{v_N^{j,p}}^{0,N_j^p} \left\{ M_j^{N,p} \in \mathcal{K}_j^p \right\} \right). \quad (120)$$

Therefore, by Assumption 6.1 we get

$$\begin{aligned}
 \limsup_{N \rightarrow \infty} \frac{1}{N} \log P_{\nu_N}^{0,N} \left(\prod_{j=1}^r \mathcal{K}_j^c \times \mathcal{K}_j^p \right) &= \limsup_{N \rightarrow \infty} \frac{1}{N} \log \left(\prod_{j=1}^r P_{\nu_N^{j,c}}^{0,N_j^c}(\mathcal{K}_j^c) P_{\nu_N^{j,p}}^{0,N_j^p}(\mathcal{K}_j^p) \right) \\
 &= \limsup_{N \rightarrow \infty} \sum_{j=1}^r \left(\frac{N_j}{N} \frac{N_j^c}{N_j} \frac{1}{N_j^c} \log P_{\nu_N^{j,c}}^{0,N_j^c}(\mathcal{K}_j^c) \right. \\
 &\qquad \qquad \qquad \left. + \frac{N_j}{N} \frac{N_j^p}{N_j} \frac{1}{N_j^p} \log P_{\nu_N^{j,p}}^{0,N_j^p}(\mathcal{K}_j^p) \right) \\
 &\leq \sum_{j=1}^r \left(\alpha_j p_j^c \limsup_{N_j^c \rightarrow \infty} \frac{1}{N_j^c} \log P_{\nu_N^{j,c}}^{0,N_j^c}(\mathcal{K}_j^c) \right. \\
 &\qquad \qquad \qquad \left. + \alpha_j p_j^p \limsup_{N_j^p \rightarrow \infty} \frac{1}{N_j^p} \log P_{\nu_N^{j,p}}^{0,N_j^p}(\mathcal{K}_j^p) \right) \\
 &\leq \sum_{j=1}^r \left(-\alpha_j p_j^c \inf_{Q_j^c \in \mathcal{K}_j^c} J^{j,c}(Q_j^c) - \alpha_j p_j^p \inf_{Q_j^p \in \mathcal{K}_j^p} J^{j,p}(Q_j^p) \right) \\
 &= - \inf_{\substack{Q_j^c \in \mathcal{K}_1^c \\ Q_j^p \in \mathcal{K}_1^p \\ \vdots \\ Q_r^c \in \mathcal{K}_r^c \\ Q_r^p \in \mathcal{K}_r^p}} \sum_{j=1}^r \left(\alpha_j p_j^c J^{j,c}(Q_j^c) + \alpha_j p_j^p J^{j,p}(Q_j^p) \right). \tag{121}
 \end{aligned}$$

Similar arguments allow us to prove the lower bound for the large deviations principle, which concludes the proof. □

The next result gives a characterization of the space containing the probability measures satisfying $L(\mathbf{Q}) < \infty$.

Lemma 6.2. *If, for a given $\mathbf{Q} = (Q_1^c, Q_1^p, \dots, Q_r^c, Q_r^p) \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$, the action functional $L(\mathbf{Q}) < \infty$, then the following hold:*

1. $\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$.
2. $\mathbf{Q} \circ \pi_0^{-1} = \nu$. Thus,

$$\left(\pi_0^{-1}(Q_1^c), \pi_0^{-1}(Q_1^p), \dots, \pi_0^{-1}(Q_r^c), \pi_0^{-1}(Q_r^p) \right) = \left(\nu^{1,c}, \nu^{1,p}, \dots, \nu^{r,c}, \nu^{r,p} \right).$$

Proof. This is a generalization of [12, Lemma 5.2]. Recall that the function $\varphi(x) = \sum_{0 \leq t \leq T} \mathbb{1}_{x(t-) \neq x(t)}$ denotes the number of jumps of x in the interval $[0, T]$. From (112) we have that $\|\varphi\|_\varphi \leq 1$. Moreover, φ is continuous in the topology induced by the metric d defined in (111). Hence $\varphi \in C_\varphi(\mathcal{X})$. Furthermore, $L(\mathbf{Q}) < \infty$ implies that, for all $1 \leq j \leq r$,

$$\int_{\mathcal{X}} \varphi dQ_j^c - \sum_{z_0 \in \mathcal{Z}} \nu^{j,c}(z) \log \int_{\mathcal{X}} e^\varphi dP_{z_0} < \infty \tag{122}$$

and

$$\int_{\mathcal{X}} \varphi dQ_j^p - \sum_{z_0 \in \mathcal{Z}} \nu^{j,p}(z) \log \int_{\mathcal{X}} e^\varphi dP_{z_0} < \infty. \tag{123}$$

Now, note that under the non-interacting distribution P_{z_0} , the transition rates are bounded by 1. Since the number of allowed transitions from any state is at most equal to $K - 1$, φ is thus stochastically dominated by a Poisson random variable of rate $(K - 1)T$. Therefore, for any initial condition $z_0 \in \mathcal{Z}$, we have $1 \leq \int_{\mathcal{X}} e^\varphi dP_{z_0} < \infty$. It follows from (122) and (123) that $\int_{\mathcal{X}} \varphi dQ_j^c < \infty$ and $\int_{\mathcal{X}} \varphi dQ_j^p < \infty$ for each $1 \leq j \leq r$, and so $\mathbf{Q} = (Q_1^c, Q_1^p, \dots, Q_r^c, Q_r^p) \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, which proves the first claim. In order to prove the second point, we proceed by contraposition. Suppose that for a given measure \mathbf{Q} , $L(\mathbf{Q}) < \infty$ and $\mathbf{Q} \circ \pi_0^{-1} = \nu_{\mathbf{Q}} \neq \nu$. Consider the bounded continuous functions $f_1^c(x), f_1^p(x), \dots, f_r^c(x), f_r^p(x)$ defined on \mathcal{X} and depending on x only through the initial condition; that is, there exist functions $g_1^c, g_1^p, \dots, g_r^c, g_r^p$ such that, for all $1 \leq j \leq r$,

$$f_j^c(x) = g_j^c(\pi_0(x)) \quad \text{and} \quad f_j^p(x) = g_j^p(\pi_0(x)).$$

Since $\nu_{\mathbf{Q}} \neq \nu$, the above functions satisfy the following claim: either

$$\sum_z g_j^c(z) \nu_{\mathbf{Q}}^{j,c}(z) - \sum_z g_j^c(z) \nu^{j,c}(z) \neq 0 \tag{124}$$

or

$$\sum_z g_j^p(z) \nu_{\mathbf{Q}}^{j,p}(z) - \sum_z g_j^p(z) \nu^{j,p}(z) \neq 0 \tag{125}$$

for at least one $1 \leq j \leq r$. Therefore, one can always find, for at least one j , an arbitrarily large $a_j^c > 0$ (or $a_j^p > 0$) such that $\sum_z g_j^c(z) \nu_{\mathbf{Q}}^{j,c}(z) - \sum_z g_j^c(z) \nu^{j,c}(z) = a_j^c$ (or $\sum_z g_j^p(z) \nu_{\mathbf{Q}}^{j,p}(z) - \sum_z g_j^p(z) \nu^{j,p}(z) = a_j^p$). Indeed, this can be done by flipping the sign of f_j^c (or f_j^p) if necessary and scaling the functions. Note that, by the assumption, $f_j^c, f_j^p \in C_\varphi(\mathcal{X})$ since they are bounded and continuous. Suppose without loss of generality that, for a given j , (124) is satisfied; then by direct calculation we obtain

$$\begin{aligned} \int_{\mathcal{X}} f_j^c dQ_j^c - \sum_{z_0 \in \mathcal{Z}} \nu^{j,c}(z) \log \int_{\mathcal{X}} e^{f_j^c} dP_{z_0} &= \int_{\mathcal{X}} g_j^c(\pi_0(x)) Q_j^c(dx) \\ &\quad - \sum_{z_0 \in \mathcal{Z}} \nu^{j,c}(z) \log \int_{\mathcal{X}} \exp \{g_j^c(\pi_0(x))\} dP_{z_0} \\ &= \sum_{z_0} g_j^c(z_0) \nu_{\mathbf{Q}}^{j,c}(z_0) - \sum_{z_0} g_j^c(z) \nu^{j,c}(z_0) = a. \end{aligned}$$

Hence, since $a > 0$ may be arbitrarily large, one gets that $J(Q_j^c) = \infty$ and then $L(\mathbf{Q}) = \infty$, which contradicts the assumption of the lemma and thus proves the second claim. \square

Conditions for application of the Laplace–Varadhan lemma. Lemma 6.1 establishes the large deviations principle for the sequence $(P_{\nu_N}^{0,N}, N \geq 1)$ in the topological space $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$. Moreover, recall that the Radon–Nikodym derivative is given by

$$\frac{dP_{\nu_N}^N}{dP_{\nu_N}^{0,N}}(\mathbf{Q}) = \exp \{Nh^N(\mathbf{Q})\}, \tag{126}$$

where the function $h^N(\mathbf{Q})$ is given by (109). Therefore, to find the large deviations principle for $(P_{v_N}^N, N \geq 1)$, one can apply the Laplace–Varadhan principle (cf. [64, Theorem 3.4]) to the sequence $(P_{v_N}^{0,N}, N \geq 1)$. The Laplace–Varadhan principle holds under the following conditions:

1. The sequence of functions $\{h^N(\mathbf{Q})\}$ satisfies

$$\lim_{A \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{Nh^N(\mathbf{Q})\} dP_{v_N}^{0,N} = -\infty. \tag{127}$$

2. For every \mathbf{Q} in $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ such that $L(\mathbf{Q}) < \infty$ and $\mathbf{Q}^N \rightarrow \mathbf{Q}$,

$$\limsup_N h^N(\mathbf{Q}^N) \leq h(\mathbf{Q}). \tag{128}$$

3. For every \mathbf{Q} in $(\mathcal{M}_{1,\varphi}(\mathcal{X}))_0^{2r}$ and $\mathbf{Q}^N \rightarrow \mathbf{Q}$,

$$\liminf_N h^N(\mathbf{Q}^N) \geq h(\mathbf{Q}), \tag{129}$$

where $(\mathcal{M}_{1,\varphi}(\mathcal{X}))_0^{2r}$ is the set of points \mathbf{Q}^* in $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ for which, given any $\varepsilon > 0$, there exists a neighborhood V of \mathbf{Q}^* such that, for any $\mathbf{Q} \in V$ and N large enough,

$$h^N(\mathbf{Q}) \geq h(\mathbf{Q}^*) - \varepsilon. \tag{130}$$

4. We have

$$\sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} [h(\mathbf{Q}) - L(\mathbf{Q})] = \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))_0^{2r}} [h(\mathbf{Q}) - L(\mathbf{Q})]. \tag{131}$$

Note that, on the one hand, the condition in (127) is satisfied if, for any $\alpha > 0$,

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \{N\alpha|h^N|\} dP_{v_N}^{0,N} < \infty. \tag{132}$$

Indeed, take $\alpha > 1$; then

$$\begin{aligned} \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{Nh^N(\mathbf{Q})\} dP_{v_N}^{0,N} &= A + \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{N(h^N(\mathbf{Q}) - A)\} dP_{v_N}^{0,N} \\ &\leq A + \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{\alpha N(h^N(\mathbf{Q}) - A)\} dP_{v_N}^{0,N} \\ &= (1 - \alpha)A + \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{\alpha Nh^N(\mathbf{Q})\} dP_{v_N}^{0,N} \\ &\leq (1 - \alpha)A + \frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \{\alpha N|h^N(\mathbf{Q})|\} dP_{v_N}^{0,N}. \end{aligned}$$

Therefore,

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{h^N(\mathbf{Q}) \geq A} \exp \{Nh^N(\mathbf{Q})\} dP_{v_N}^{0,N} \leq \limsup_{N \rightarrow \infty} \left((1 - \alpha)A + \frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \{\alpha N|h^N(\mathbf{Q})|\} dP_{v_N}^{0,N} \right),$$

where the right-hand side of the last inequality goes to $-\infty$ as $A \rightarrow \infty$ if (132) is true. On the other hand, it is easy to see that the conditions (128), (129), and (131) hold if the functions h^N defined in (109) are continuous and the sequence $\{h^N\}$ converges uniformly on $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ towards the function h given in (110). The next lemmas are thus dedicated to the verification of these conditions. First, we establish a regularity property for all the probability measures \mathbf{Q} satisfying $L(\mathbf{Q}) < \infty$.

Lemma 6.3. *Let $\mathbf{Q} = (Q_1^c, Q_1^p, \dots, Q_r^c, Q_r^p) \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ be such that $L(\mathbf{Q}) < \infty$. Moreover, suppose that the random vector $\mathbf{X} = (X_1^c, X_1^p, \dots, X_r^c, X_r^p)$ is distributed according to \mathbf{Q} . Then*

$$\sup_{t \in [0, T]} \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{\mathbf{X}(u) \neq \mathbf{X}(u-)} \} \right] \rightarrow 0 \quad \text{as } \alpha \downarrow 0. \tag{133}$$

Proof. This is a generalization of [12, Lemma 5.7]. Note that $\mathbf{X}(u) \neq \mathbf{X}(u-)$ if $X_j^c(u) \neq X_j^c(u-)$ or $X_j^p(u) \neq X_j^p(u-)$ for at least one $1 \leq j \leq r$. Therefore, for each $t \in [0, T]$ one obtains

$$\begin{aligned} \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{\mathbf{X}(u) \neq \mathbf{X}(u-)} \} \right] &\leq \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_1^c(u) \neq X_1^c(u-)} \} \right] \\ &+ \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_1^p(u) \neq X_1^p(u-)} \} \right] \\ &+ \dots + \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_r^c(u) \neq X_r^c(u-)} \} \right] + \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_r^p(u) \neq X_r^p(u-)} \} \right]. \end{aligned} \tag{134}$$

Moreover, since $L(\mathbf{Q}) < \infty$, one gets that $J(Q_j^c) < \infty$ and $J(Q_j^p) < \infty$ for all $1 \leq j \leq r$. Hence, applying [12, Lemma 5.7] to each of the X_j^c and X_j^p with respective marginal distributions Q_j^c and Q_j^p gives us that, for each $1 \leq j \leq r$,

$$\sup_{t \in [0, T]} \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_j^c(u) \neq X_j^c(u-)} \} \right] \rightarrow 0 \quad \text{as } \alpha \downarrow 0 \tag{135}$$

and

$$\sup_{t \in [0, T]} \mathbb{E} \left[\sup_{u \in [t-\alpha, t+\alpha] \cap [0, T]} \{ \mathbb{1}_{X_j^p(u) \neq X_j^p(u-)} \} \right] \rightarrow 0 \quad \text{as } \alpha \downarrow 0. \tag{136}$$

Combining (134), (135), and (136) leads to (133). □

The next lemma establishes the continuity of the projection π , which is needed for the continuity of the function $h(\mathbf{Q})$.

Lemma 6.4. *Let $\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}))$ be equipped with its usual weak topology and let $\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))$ be equipped with the metric*

$$\rho_T(\mu, \nu) = \sup_{0 \leq t \leq T} \rho_0(\mu_t, \nu_t), \quad \mu, \nu \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})), \tag{137}$$

where $\rho_0(\cdot, \cdot)$ is a metric on $\mathcal{M}_1(\mathcal{Z})$ which generates the weak topology $\sigma(\mathcal{M}_1(\mathcal{Z}), C_b(\mathcal{Z}))$. Moreover, let $(\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ be endowed with the product topology induced by the product metric. Equivalently, let $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$ be equipped with the product topology obtained from the product metric $\rho_T^{2r} = \max\{\rho_T, \dots, \rho_T\}$. Then the projection

$$\pi : \mathbf{Q} \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r} \rightarrow \pi(\mathbf{Q}) = (\mathbf{Q}_t)_{0 \leq t \leq T} \in (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$$

is continuous at each $\mathbf{Q} \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ where $L(\mathbf{Q}) < \infty$.

Proof. The statement of our lemma resembles the statement of [49, Lemma 2.8]. The difference here is that our spaces of interest are the product spaces $(\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ and $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$ endowed with the product metrics. Moreover, the rate $J(\mathbf{Q})$ in [49, Lemma 2.8] is here replaced by $L(\mathbf{Q})$. Therefore, if we replace the norm $|\cdot|$ by the product norm $\|\cdot\|$ adapted to the context of our product spaces, then the proof of our lemma follows verbatim the proof of [49, Lemma 2.8], provided that we can prove [49, Equation (2.14)]. This is done in Lemma 6.3. Thus, the proof is complete. \square

We next prove the continuity of the functions h^N .

Lemma 6.5. *The functions $h^N : (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r} \rightarrow \mathbb{R}$ defined at (109) are continuous at any \mathbf{Q} such that $L(\mathbf{Q}) < \infty$.*

Proof. This is a generalization of [49, Lemma 2.9]. For any $\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, define

$$\theta_{\mathbf{Q}}^{j,c}(x) := \sum_{0 \leq t \leq T} \mathbb{1}_{x_t \neq x_{t-}} \log \left(\sum_{(x_{t-}, x(t)) \in \mathcal{E}} \lambda_{j,x_{t-},x_t}^c \left(\mathcal{Q}_j^c(t-), \mathcal{Q}_j^p(t-) \right) \right), \tag{138}$$

$$\theta_{\mathbf{Q}}^{j,p}(x) := \sum_{0 \leq t \leq T} \mathbb{1}_{x_t \neq x_{t-}} \log \left(\sum_{(x_{t-}, x_t) \in \mathcal{E}} \lambda_{j,x_{t-},x_t}^p \left(\mathcal{Q}_j^c(t-), \mathcal{Q}_1^p(t-), \dots, \mathcal{Q}_r^p(t-) \right) \right), \tag{139}$$

$$\gamma_{\mathbf{Q}}^{j,c}(x) := \int_0^T \left(\sum_{z: (x_r, z) \in \mathcal{E}} \lambda_{j,x_r,z}^c \left(\mathcal{Q}_j^c(t), \mathcal{Q}_j^p(t) \right) - 1 \right) dt, \tag{140}$$

$$\gamma_{\mathbf{Q}}^{j,p}(x) := \int_0^T \left(\sum_{z: (x_r, z) \in \mathcal{E}} \lambda_{j,x_r,z}^p \left(\mathcal{Q}_j^c(t), \mathcal{Q}_1^p(t), \dots, \mathcal{Q}_r^p(t) \right) - 1 \right) dt. \tag{141}$$

Note that the function h^N given by (109) can be rewritten using the functions $\theta_{\mathbf{Q}}^{j,c}(x)$, $\theta_{\mathbf{Q}}^{j,p}(x)$, $\gamma_{\mathbf{Q}}^{j,c}(x)$, and $\gamma_{\mathbf{Q}}^{j,p}(x)$ as follows:

$$h^N(\mathbf{Q}) = \sum_{j=1}^r \left[\frac{N_j^c}{N} \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,c}(x) - \gamma_{\mathbf{Q}}^{j,c}(x) \right) \mathcal{Q}_j^c(dx) + \frac{N_j^p}{N} \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,p}(x) - \gamma_{\mathbf{Q}}^{j,p}(x) \right) \mathcal{Q}_j^p(dx) \right]. \tag{142}$$

Therefore, to show the continuity of $h^N(\mathbf{Q})$, we show that for any $1 \leq j \leq r$, the functions

$$\begin{aligned} \mathbf{Q} &\rightarrow \int_{\mathcal{X}} \theta_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx), & \mathbf{Q} &\rightarrow \int_{\mathcal{X}} \theta_{\mathbf{Q}}^{j,p}(x) \mathcal{Q}_j^p(dx), \\ \mathbf{Q} &\rightarrow \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx), & \mathbf{Q} &\rightarrow \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,p}(x) \mathcal{Q}_j^p(dx) \end{aligned}$$

are continuous at any \mathbf{Q} where $L(\mathbf{Q}) < \infty$. First, from Assumption 6.1, there exists a positive constant $C > 0$ such that, for each $1 \leq j \leq r$,

$$\begin{aligned} \left| \theta_{\mathbf{Q}}^{j,c}(x) \right| &\leq \sup_{\xi, \zeta} \left(\left| \log \left(\sum_{(x_t-, x_t) \in \mathcal{E}} \lambda_{j, x_t-, x_t}^c(\xi, \zeta) \right) \right| \right) \varphi(x) \\ &\leq C(1 + \varphi(x)) \quad \forall x \in \mathcal{X} \end{aligned} \tag{143}$$

and

$$\begin{aligned} \left| \theta_{\mathbf{Q}}^{j,p}(x) \right| &\leq \sup_{\xi, \zeta_1, \dots, \zeta_r} \left(\left| \log \left(\sum_{(x_t-, x_t) \in \mathcal{E}} \lambda_{j, x_t-, x_t}^p(\xi, \zeta_1, \dots, \zeta_r) \right) \right| \right) \varphi(x) \\ &\leq C(1 + \varphi(x)) \quad \forall x \in \mathcal{X}. \end{aligned} \tag{144}$$

Similarly, by Assumption 6.1 we have that, for each $1 \leq j \leq r$,

$$\left| \gamma_{\mathbf{Q}}^{j,c}(x) \right| \leq \sup_{\xi, \zeta} \left| \int_0^T \left(\sum_{z: (x_t, z) \in \mathcal{E}} \lambda_{j, x_t, z}^c(\xi, \zeta) - 1 \right) dt \right| < \infty \quad \forall x \in \mathcal{X} \tag{145}$$

and

$$\left| \gamma_{\mathbf{Q}}^{j,p}(x) \right| \leq \sup_{\xi, \zeta_1, \dots, \zeta_r} \left| \int_0^T \left(\sum_{z: (x_t, z) \in \mathcal{E}} \lambda_{j, x_t, z}^p(\xi, \zeta_1, \dots, \zeta_r) - 1 \right) dt \right| < \infty \quad \forall x \in \mathcal{X}. \tag{146}$$

Take $\mathbf{Q}' \in (\mathcal{M}_{1, \varphi}(\mathcal{X}))^{2r}$ in the neighborhood of \mathbf{Q} . Note that

$$\begin{aligned} |h^N(\mathbf{Q}) - h^N(\mathbf{Q}')| &\leq \sum_{j=1}^r \left[\frac{N_j^c}{N} \left(\left| \langle \theta_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c \rangle - \langle \theta_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right| + \left| \langle \gamma_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c \rangle - \langle \gamma_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right| \right) \right. \\ &\quad \left. + \frac{N_j^p}{N} \left(\left| \langle \theta_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p \rangle - \langle \theta_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right| + \left| \langle \gamma_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p \rangle - \langle \gamma_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right| \right) \right]. \end{aligned} \tag{147}$$

In addition, for each $1 \leq j \leq r$, the following inequalities hold:

$$\left| \langle \theta_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c \rangle - \langle \theta_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right| \leq \left| \langle \theta_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c - \mathcal{Q}_j'^c \rangle \right| + \left| \langle \theta_{\mathbf{Q}}^{j,c} - \theta_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right|, \tag{148}$$

$$\left| \langle \theta_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p \rangle - \langle \theta_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right| \leq \left| \langle \theta_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p - \mathcal{Q}_j'^p \rangle \right| + \left| \langle \theta_{\mathbf{Q}}^{j,p} - \theta_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right|, \tag{149}$$

$$\left| \langle \gamma_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c \rangle - \langle \gamma_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right| \leq \left| \langle \gamma_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c - \mathcal{Q}_j'^c \rangle \right| + \left| \langle \gamma_{\mathbf{Q}}^{j,c} - \gamma_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right|, \tag{150}$$

$$\left| \langle \gamma_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p \rangle - \langle \gamma_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right| \leq \left| \langle \gamma_{\mathbf{Q}}^{j,p}, \mathcal{Q}_j^p - \mathcal{Q}_j'^p \rangle \right| + \left| \langle \gamma_{\mathbf{Q}}^{j,p} - \gamma_{\mathbf{Q}'}^{j,p}, \mathcal{Q}_j'^p \rangle \right|. \tag{151}$$

The idea now is to control the right-hand sides of the last four inequalities. We show this for the inequality in (148). Similar arguments can be used to treat the other three inequalities. First, notice that the function $\theta_{\mathbf{Q}}^{j,c}$ is continuous. Indeed, the topology of \mathcal{X} is built so that the function $x \rightarrow \sum_{0 \leq t \leq T} \mathbb{1}_{x_t \neq x_{t-}}$ is continuous. Moreover, from Assumption 6.1, the functions $\lambda_{j,z,z}^c$ are continuous. Furthermore, from Lemma 6.4, the component projection $\mathcal{Q}_j^c \rightarrow \pi(\mathcal{Q}_j^c) = (\mathcal{Q}_j^c(t))_{0 \leq t \leq T}$ is continuous since $\pi(\mathbf{Q}) = (\mathbf{Q}(t))_{0 \leq t \leq T}$ is continuous. Finally, the log function being continuous gives that $\theta_{\mathbf{Q}}^{j,c}$ is continuous provided that $L(\mathbf{Q}) < \infty$. In addition, from (143) we have that $\theta_{\mathbf{Q}}^{j,c} \leq C(1 + \varphi)$; thus $\theta_{\mathbf{Q}}^{j,c} \in C_\varphi(\mathcal{X})$ provided that $L(\mathbf{Q}) < \infty$. Therefore, the term $\left| \langle \theta_{\mathbf{Q}}^{j,c}, \mathcal{Q}_j^c - \mathcal{Q}_j'^c \rangle \right|$ can be made as small as desired by taking \mathbf{Q}' close enough to \mathbf{Q} (and thus $\mathcal{Q}_j'^c$ close enough to \mathcal{Q}_j^c). The second term in the right-hand side of (148) is bounded as follows:

$$\begin{aligned} \left| \langle \theta_{\mathbf{Q}}^{j,c} - \theta_{\mathbf{Q}'}^{j,c}, \mathcal{Q}_j'^c \rangle \right| &\leq \sup_t \left| \log \left(\sum_{(x_{t-}, x_t) \in \mathcal{E}} \lambda_{j,x_{t-},x_t}^c \left(\mathcal{Q}_j^c(t-), \mathcal{Q}_j^p(t-) \right) \right) \right. \\ &\quad \left. - \log \left(\sum_{(x_{t-}, x_t) \in \mathcal{E}} \lambda_{j,x_{t-},x_t}^c \left(\mathcal{Q}_j'^c(t-), \mathcal{Q}_j'^p(t-) \right) \right) \right| \int_{\mathcal{X}} \varphi d\mathcal{Q}_j'^c. \end{aligned} \tag{152}$$

Therefore, using again Assumption 6.1, Lemma 6.4, and the continuity of the log function, the right-hand side of (152) is controlled for any \mathbf{Q}' in the neighborhood of \mathbf{Q} in $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ provided that $L(\mathbf{Q}) < \infty$. Thus, the integral $\mathbf{Q} \rightarrow \int \theta_{\mathbf{Q}}^{j,c} d\mathcal{Q}_j^c$ is continuous. The same steps allow us to show that

$$\mathbf{Q} \rightarrow \int_{\mathcal{X}} \theta_{\mathbf{Q}}^{j,p}(x) \mathcal{Q}_j^p(dx), \quad \mathbf{Q} \rightarrow \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx), \quad \text{and} \quad \mathbf{Q} \rightarrow \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,p}(x) \mathcal{Q}_j^p(dx)$$

are also continuous at any \mathbf{Q} where $L(\mathbf{Q}) < \infty$. Hence, the function h^N is a linear combination of continuous functions and thus is continuous, which concludes the proof. \square

The following lemma states the uniform convergence of $\{h^N, N \geq 1\}$ towards h .

Lemma 6.6. *Suppose Assumption 6.1 holds. Then the sequence of functions $\{h^N, N \geq 1\}$ introduced in (109) converges uniformly towards the function h given by (110).*

Proof. From (110) and (142) one has

$$\begin{aligned} \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} |h^N(\mathbf{Q}) - h(\mathbf{Q})| &= \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \sum_{j=1}^r \left[\left(\frac{N_j^c}{N} - \alpha_j p_j^c \right) \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,c}(x) - \gamma_{\mathbf{Q}}^{j,c}(x) \right) \mathcal{Q}_j^c(dx) \right. \right. \\ &\quad \left. \left. + \left(\frac{N_j^p}{N} - \alpha_j p_j^p \right) \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,p}(x) - \gamma_{\mathbf{Q}}^{j,p}(x) \right) \mathcal{Q}_j^p(dx) \right] \right| \end{aligned}$$

$$\begin{aligned} &\leq \sum_{j=1}^r \left(\left| \frac{N_j^c}{N} - \alpha_j p_j^c \right| \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,c}(x) - \gamma_{\mathbf{Q}}^{j,c}(x) \right) \mathcal{Q}_j^c(dx) \right| \right. \\ &\quad \left. + \left| \frac{N_j^p}{N} - \alpha_j p_j^p \right| \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,p}(x) - \gamma_{\mathbf{Q}}^{j,p}(x) \right) \mathcal{Q}_j^p(dx) \right| \right). \end{aligned} \tag{153}$$

First, observe that

$$\begin{aligned} \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,c}(x) - \gamma_{\mathbf{Q}}^{j,c}(x) \right) \mathcal{Q}_j^c(dx) \right| &\leq \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \theta_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx) \right| \\ &\quad + \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx) \right|. \end{aligned}$$

Using (143) one obtains

$$\begin{aligned} \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \theta_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx) \right| &\leq \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \int_{\mathcal{X}} C(1 + \varphi(x)) \mathcal{Q}_j^c(dx) \\ &\leq \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} C \left(\int_{\mathcal{X}} \mathcal{Q}_j^c(dx) + \int_{\mathcal{X}} \varphi(x) \mathcal{Q}_j^c(dx) \right), \end{aligned}$$

which is $< \infty$ since \mathcal{Q}_j^c is a probability measure and the second integral is finite for $\mathcal{Q}_j^c \in \mathcal{M}_{1,\varphi}(\mathcal{X})$. Moreover, by (145) one has

$$\sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \gamma_{\mathbf{Q}}^{j,c}(x) \mathcal{Q}_j^c(dx) \right| < \infty,$$

again since \mathcal{Q}_j^c is a probability measure. Therefore, by Assumption 6.1, one deduces that

$$\left(\frac{N_j^c}{N} - \alpha_j p_j^c \right) \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,c}(x) - \gamma_{\mathbf{Q}}^{j,c}(x) \right) \mathcal{Q}_j^c(dx) \right| \xrightarrow{N \rightarrow \infty} 0.$$

Similarly, one obtains

$$\left(\frac{N_j^p}{N} - \alpha_j p_j^p \right) \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \left| \int_{\mathcal{X}} \left(\theta_{\mathbf{Q}}^{j,p}(x) - \gamma_{\mathbf{Q}}^{j,p}(x) \right) \mathcal{Q}_j^p(dx) \right| \xrightarrow{N \rightarrow \infty} 0.$$

Thus h^N converges uniformly towards h . □

The final step before applying the Laplace–Varadhan principle is to verify that (132) is satisfied.

Lemma 6.7. *For any $\alpha > 0$,*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \{ N\alpha |h^N| \} dP_{v_N}^{0,N} < \infty.$$

Proof. First note that, using the bounds (143), (144), (145), and (146), we find that for all $\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$,

$$|h^N(\mathbf{Q})| \leq \sum_{j=1}^r \left[\frac{N_j^c}{N} C \left(1 + \int_{\mathcal{X}} \varphi(x) Q_j^c(dx) \right) + \frac{N_j^p}{N} C \left(1 + \int_{\mathcal{X}} \varphi(x) Q_j^p(dx) \right) \right]. \tag{154}$$

Therefore, to show (132), it is enough to show that, for any $\alpha > 0$,

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \left\{ N\alpha \sum_{j=1}^r \left[\frac{N_j^c}{N} \int_{\mathcal{X}} \varphi(x) Q_j^c(dx) + \frac{N_j^p}{N} \int_{\mathcal{X}} \varphi(x) Q_j^p(dx) \right] \right\} dP_{v_N}^{0,N}(\mathbf{Q}) < \infty. \tag{155}$$

Recall that $P_{v_N}^{0,N} = \mathbb{P}_{z_N}^{0,N} \circ G_N^{-1}$, where $\mathbb{P}_{z_N}^{0,N} = \otimes_{n=1}^N P_{z_n}$ and P_{z_n} is the law of the n th particle in the case of non-interaction, with the initial condition being z_n . Hence, by independence, the integral term in the left-hand side of (155) is equivalent to

$$\prod_{j=1}^r \left(\int_{\mathcal{M}_{1,\varphi}(\mathcal{X})} \exp \left\{ N_j^c \alpha \int_{\mathcal{X}} \varphi(x) Q_j^c(dx) \right\} dP_{v_N^{j,c}}^{0,N_j^c}(Q_j^c) \int_{\mathcal{M}_{1,\varphi}(\mathcal{X})} \exp \left\{ N_j^p \alpha \int_{\mathcal{X}} \varphi(x) Q_j^p(dx) \right\} dP_{v_N^{j,p}}^{0,N_j^p}(Q_j^p) \right). \tag{156}$$

Now, using [49, Lemma 2.10], we find that for all $1 \leq j \leq r$,

$$\limsup_{N_j^c \rightarrow \infty} \frac{1}{N_j^c} \log \int_{\mathcal{M}_{1,\varphi}(\mathcal{X})} \exp \left\{ N_j^c \alpha \int_{\mathcal{X}} \varphi(x) Q_j^c(dx) \right\} dP_{v_N^{j,c}}^{0,N_j^c}(Q_j^c) < \infty \tag{157}$$

and

$$\limsup_{N_j^p \rightarrow \infty} \frac{1}{N_j^p} \log \int_{\mathcal{M}_{1,\varphi}(\mathcal{X})} \exp \left\{ N_j^p \alpha \int_{\mathcal{X}} \varphi(x) Q_j^p(dx) \right\} dP_{v_N^{j,p}}^{0,N_j^p}(Q_j^p) < \infty. \tag{158}$$

Since $N_j^c < N$ and $N_j^p < N$ for all $1 \leq j \leq r$, (157), (158), and (156) lead to (155), which concludes the proof. \square

The interacting case. We are now ready to apply the Laplace–Varadhan principle to the sequence of probability measures $\{P_{v_N}^{0,N}, N \geq 1\}$. By Lemma 6.1, the sequence $\{P_{v_N}^{0,N}, N \geq 1\}$ obeys a large deviations principle in the topological space $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$ with rate function $L(\mathbf{Q})$ defined by (117), and with speed N . Moreover, by Lemma 6.5, the functions h^N defined in (109) are continuous at any \mathbf{Q} such that $L(\mathbf{Q}) < \infty$. Furthermore, by Lemma 6.2, the functions h^N are continuous on the set $\{\mathbf{Q} \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z}))^{2r} | L(\mathbf{Q}) < \infty\}$. Therefore, the conditions in (128), (129), and (131) hold true. Finally, we have seen in Lemma 6.7 that (132) is satisfied. Hence, a straightforward application of [64, Theorem 3.4] gives

$$\frac{1}{N} \log \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp \{Nh^N\} dP_{v_N}^{0,N} \longrightarrow \sup_{\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} [h(\mathbf{Q}) - L(\mathbf{Q})] \tag{159}$$

as $N \rightarrow \infty$, and the sequence

$$\left\{ \frac{\exp(Nh^N)}{\int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp(Nh^N) dP_{v^N}^{0,N}} \cdot P_{v^N}^{0,N}, N \geq 1 \right\} \tag{160}$$

obeys a large deviations principle with speed N and rate function

$$\mathbf{Q} \rightarrow L(\mathbf{Q}) - h(\mathbf{Q}) - \inf_{\mathbf{Q}' \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} [L(\mathbf{Q}') - h(\mathbf{Q}')]. \tag{161}$$

Now, from (108) we have

$$\frac{dP_{v^N}^N}{dP_{v^N}^{0,N}}(\mathbf{Q}) = \exp\{Nh^N(\mathbf{Q})\}. \tag{162}$$

Since $P_{v^N}^N$ is a probability measure we obtain

$$\int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} \exp(Nh^N) dP_{v^N}^{0,N} = \int_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} dP_{v^N}^N = 1. \tag{163}$$

Thus, the left-hand side of (159) is always zero and so

$$\sup_{(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} [h(\mathbf{Q}) - L(\mathbf{Q})] = 0, \tag{164}$$

which gives that

$$\inf_{\mathbf{Q}' \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}} [L(\mathbf{Q}') - h(\mathbf{Q}')] = 0. \tag{165}$$

We then conclude that the sequence $\{P_{v^N}^N, N \geq 1\}$ obeys a large deviations principle in the topological space $(\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, with speed N and rate function

$$I(\mathbf{Q}) = L(\mathbf{Q}) - h(\mathbf{Q}). \tag{166}$$

In order to obtain the representation in (119), we proceed as follows. First, from (154) we have that, for $\mathbf{Q} \in (\mathcal{M}_{1,\varphi}(\mathcal{X}))^{2r}$, $h(\mathbf{Q}) < \infty$. Moreover, from [12, Lemma 5.6], the functions $J^{j,\iota}(Q)$ defined by (118) have the following representation:

$$J^{j,\iota}(Q) = \begin{cases} H(Q|P_{j,\iota}) & \text{if } Q \circ \pi_0^{-1} = v^{j,\iota}, \\ +\infty & \text{otherwise,} \end{cases} \tag{167}$$

where $H(Q|P_{j,\iota})$ is the relative entropy defined by (116). Therefore, if either $Q \circ \pi_0^{-1} \neq v^{j,\iota}$ or Q is not absolutely continuous with respect to $P_{j,\iota}$, then one can immediately observe that $J^{j,\iota}(Q) = \infty$; thus $L(\mathbf{Q}) = \infty$, and finally $I(\mathbf{Q}) = \infty$. Now, assume that for all $1 \leq j \leq r$ we have $Q_j^c \circ \pi_0^{-1} = v^{j,c}$, $Q_j^c \ll P$, and $Q_j^p \circ \pi_0^{-1} = v^{j,p}$, $Q_j^p \ll P$; then

$$L(\mathbf{Q}) = \alpha_1 p_1^c H(Q_1^c|P_{1,c}) + \alpha_1 p_1^p H(Q_1^p|P_{1,p}) + \dots + \alpha_r p_r^c H(Q_r^c|P_{r,c}) + \alpha_r p_r^p H(Q_r^p|P_{r,p}).$$

Furthermore, one can observe from (105) that the densities $\exp \{h_1(x, \eta, \rho_j)\}$ and $\exp \{h_2(x, \eta, \rho_1, \dots, \rho_r)\}$ do not depend on the initial condition z_0 . Therefore, for each $1 \leq j \leq r$, the densities of $R^c(\pi(Q_j^c), \pi(Q_j^p))$ and $R^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p))$ with respect to the mixtures $P_{j,c}$ and $P_{j,p}$ are given by, respectively,

$$\frac{dR^c(\pi(Q_j^c), \pi(Q_j^p))}{dP_{j,c}}(x) = \exp \left\{ h_1(x, \pi(Q_j^c), \pi(Q_j^p)) \right\}$$

and

$$\frac{dR^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p))}{dP_{j,p}}(x) = \exp \left\{ h_2(x, \pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p)) \right\}.$$

Replacing $h_1(\cdot)$ and $h_2(\cdot)$ in (109) by the two representations above, we find

$$h(\mathbf{Q}) = \sum_{j=1}^r \left[\frac{N_j^c}{N} \int_{D([0,T], \mathcal{Z})} dQ_j^c \log \frac{dR^c(\pi(Q_j^c), \pi(Q_j^p))}{dP_{j,c}} + \frac{N_j^p}{N} \int_{D([0,T], \mathcal{Z})} dQ_j^p \log \frac{dR^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p))}{dP_{j,p}} \right]. \tag{168}$$

Finally, using (99) we find that, as $N \rightarrow \infty$,

$$\begin{aligned} L(\mathbf{Q}) - h(\mathbf{Q}) &= \sum_{j=1}^r \left[\alpha_j p_j^c \left(\int_{D([0,T], \mathcal{Z})} dQ_j^c \log \frac{dQ_j^c}{dP_{j,c}} - \int_{D([0,T], \mathcal{Z})} dQ_j^c \log \frac{dR^c(\pi(Q_j^c), \pi(Q_j^p))}{dP_{j,c}} \right) \right. \\ &\quad \left. + \alpha_j p_j^p \left(\int_{D([0,T], \mathcal{Z})} dQ_j^p \log \frac{dQ_j^p}{dP} \right. \right. \\ &\quad \left. \left. - \int_{D([0,T], \mathcal{Z})} dQ_j^p \log \frac{dR^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p))}{dP} \right) \right] \\ &= \sum_{j=1}^r \left[\alpha_j p_j^c \int_{D([0,T], \mathcal{Z})} dQ_j^c \log \frac{dQ_j^c}{dR^c(\pi(Q_j^c), \pi(Q_j^p))} \right. \\ &\quad \left. + \alpha_j p_j^p \int_{D([0,T], \mathcal{Z})} dQ_j^p \log \frac{dQ_j^p}{dR^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p))} \right] \\ &= \sum_{j=1}^r \left[\alpha_j p_j^c H \left(Q_j^c \middle| R^c(\pi(Q_j^c), \pi(Q_j^p)) \right) \right. \\ &\quad \left. + \alpha_j p_j^p H \left(Q_j^p \middle| R^p(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p)) \right) \right]. \end{aligned} \tag{169}$$

This concludes the proof. □

6.2. Large deviations principle for the empirical process

We now investigate the large deviations of the sequence $(p_{v_N}^N, N \geq 1)$ where, for any $N \geq 1$, $p_{v_N}^N = \mathbb{P}_{z_N}^N \circ \gamma_N^{-1} = \pi(\mathbb{M}^N)$ is the distribution of the $(\mathcal{M}_1(\mathcal{Z}))^{2r}$ -valued empirical process

defined by

$$\begin{aligned} \mu^N : t \in [0, T] \longrightarrow \mu^N(t) &= \left(\mu_1^{c,N}(t), \mu_1^{p,N}(t), \dots, \mu_r^{c,N}(t), \mu_r^{p,N}(t) \right) \\ &= \left(\frac{1}{N_1^c} \sum_{n \in C_1^c} \delta_{X_{n,1}^c}(t), \frac{1}{N_1^p} \sum_{n \in C_1^p} \delta_{X_{n,1}^p}(t), \dots, \frac{1}{N_r^c} \sum_{n \in C_r^c} \delta_{X_{n,r}^c}(t), \frac{1}{N_r^p} \sum_{n \in C_r^p} \delta_{X_{n,r}^p}(t) \right). \end{aligned}$$

The flow μ^N takes values in the product space $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$. Again let $\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))$ be equipped with the metric

$$\rho_T(\mu, \nu) := \sup_{0 \leq t \leq T} \rho_0(\mu_t, \nu_t), \quad \mu, \nu \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})), \tag{170}$$

where $\rho_0(\alpha, \beta)$, $\alpha, \beta \in \mathcal{M}_1(\mathcal{Z})$, is a metric on $\mathcal{M}_1(\mathcal{Z})$ that generates the weak topology on $\mathcal{M}_1(\mathcal{Z})$. Moreover, let the product space $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$ be equipped with the product topology induced by the product metric $\rho_T^{2r} = \max\{\rho_T, \dots, \rho_T\}$.

For any $\xi = (\xi_1^c, \xi_1^p, \dots, \xi_r^c, \xi_r^p) \in (\mathcal{M}_1(\mathcal{Z}))^{2r}$, define the rate matrices

$$A_{\xi}^{j,c} := \left(\lambda_{j,z,z'}^c(\xi_j^c, \xi_j^p) \right)_{(z,z') \in \mathcal{Z} \times \mathcal{Z}} \quad \text{and} \quad A_{\xi}^{j,p} := \left(\lambda_{j,z,z'}^p(\xi_j^c, \xi_1^p, \dots, \xi_r^p) \right)_{(z,z') \in \mathcal{Z} \times \mathcal{Z}}, \tag{171}$$

where

$$\lambda_{j,z,z}^c(\xi_j^c, \xi_j^p) = - \sum_{z' \neq z} \lambda_{j,z,z'}^c(\xi_j^c, \xi_j^p)$$

and

$$\lambda_{j,z,z}^p(\xi_j^c, \xi_1^p, \dots, \xi_r^p) = - \sum_{z' \neq z} \lambda_{j,z,z'}^p(\xi_j^c, \xi_1^p, \dots, \xi_r^p).$$

From the laws of large numbers given in Corollary 5.1, one can deduce that, as $N \rightarrow \infty$, the sequence $(\mu^N, N \geq 1)$ converges weakly, for converging initial conditions, towards the solution μ of the following McKean–Vlasov system:

$$\begin{cases} \dot{\mu}_j^c(t) = A_{\mu(t)}^{j,c*} \mu_j^c(t), \\ \dot{\mu}_j^p(t) = A_{\mu(t)}^{j,p*} \mu_j^p(t), \\ \mu_j^c(0) = v_j^c, \mu_j^p(0) = v_j^p, \\ 1 \leq j \leq r, \end{cases} \tag{172}$$

where A^* is the adjunct/transpose of the matrix A and $\dot{\mu}(t) = \frac{\partial}{\partial t} \mu(t)$. Note that the Lipschitz property of the functions $\lambda_{j,z,z'}^c$ and $\lambda_{j,z,z'}^p$ ensures that (172) is well-posed. Also, one can notice that the representation (172) is consistent with the infinitesimal generators $\mathcal{L}_{\xi, \eta_j}^c$ and $\mathcal{L}_{\xi, \eta_1, \dots, \eta_r}^p$ introduced in (103) and (104). Indeed, if we consider ϕ , $\mathcal{L}_{\xi, \eta_j}^c \phi$, and $\mathcal{L}_{\xi, \eta_1, \dots, \eta_r}^p \phi$ as column vectors, then the right-hand sides of (103) and (104) are the results of right-multiplying the rates matrices $A^{j,c}$ and $A^{j,p}$, respectively, by the vector ϕ .

Denote by τ the log-Laplace transform of the centered Poisson distribution with parameter 1 given by $\tau(u) = e^u - u - 1$, and let τ^* be its Legendre transform, defined by

$$\tau^*(u) := \begin{cases} (u + 1) \log(u + 1) - u & \text{if } u > -1, \\ 1 & \text{if } u = -1, \\ +\infty & \text{if } u < -1. \end{cases}$$

Let us recall now the notion of absolute continuity introduced in [24, Definition 4.1]. Denote by \mathcal{S} the Schwartz space of test functions $\mathbb{R}^d \rightarrow \mathbb{R}$ having compact support and possessing continuous derivatives of all orders. We endow \mathcal{S} with the usual inductive topology. Let \mathcal{S}' be the corresponding space of real distributions. For each compact set $K \subset \mathbb{R}^d$, \mathcal{S}_K will denote the subspace of \mathcal{S} consisting of all test functions whose support is contained in K . Finally, let $\langle \nu, f \rangle$ denote the application of the test function f to the distribution ν .

Definition 6.1. Let I be an interval of the real line. A map $\nu(\cdot) : I \rightarrow \mathcal{S}'$ is called absolutely continuous if, for each compact set $K \subset \mathbb{R}^d$, there exists a neighborhood U_K of 0 in \mathcal{S}_K and an absolutely continuous function $H_K : I \rightarrow \mathbb{R}$ such that

$$|\langle \nu(u), f \rangle| - |\langle \nu(v), f \rangle| \leq H_K(u) - H_K(v)$$

for all $u, v \in I$ and $f \in U_K$.

Finally, for any $\theta \in \mathcal{M}(\mathcal{Z})$, define

$$\begin{aligned} \|\theta\|_{\mu(t)}^{j,c} &:= \sup_{\Phi : \mathcal{Z} \rightarrow \mathbb{R}} \left\{ \sum_{z \in \mathcal{Z}} \theta(z) \cdot \Phi(z) - \sum_{(z,z') \in \mathcal{E}} \tau(\Phi(z') - \Phi(z)) \cdot \mu_j^c(t)(z) \cdot \lambda_{j,z,z'}^c(\mu_j^c(t), \mu_j^p(t)) \right\}, \\ \|\theta\|_{\mu(t)}^{j,p} &:= \sup_{\Phi : \mathcal{Z} \rightarrow \mathbb{R}} \left\{ \sum_{z \in \mathcal{Z}} \theta(z) \cdot \Phi(z) \right. \\ &\quad \left. - \sum_{(z,z') \in \mathcal{E}} \tau(\Phi(z') - \Phi(z)) \cdot \mu_j^p(t)(z) \cdot \lambda_{j,z,z'}^p(\mu_j^c(t), \mu_1^p(t), \dots, \mu_r^p(t)) \right\}. \end{aligned}$$

Also let us introduce, for each $\nu \in (\mathcal{M}_1(\mathcal{Z}))^{2r}$, and according to [24, Equation (4.9)], the functional $S(\mu|\nu)$ defined from $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))^{2r})$ to $[0, \infty]$ by setting

$$\begin{aligned} S_{[0,T]}(\mu|\nu) &:= \sum_{j=1}^r \left[\alpha_j p_j^c \int_0^T \|\dot{\mu}_j^c(t) - A_{\mu(t)}^{j,c*} \mu_j^c(t)\|_{\mu(t)} dt + \alpha_j p_j^p \int_0^T \|\dot{\mu}_j^p(t) \right. \\ &\quad \left. - A_{\mu(t)}^{j,p*} \mu_j^p(t)\|_{\mu(t)} dt \right] \end{aligned} \tag{173}$$

if $\mu(0) = \nu$ and μ_j^c, μ_j^p are absolutely continuous in the sense of Definition 6.1 for all $1 \leq j \leq r$, and $S_{[0,T]}(\mu|\nu) = +\infty$ otherwise.

We are now ready to state our large deviations principle for the sequence $(p_{\nu_N}^N, N \geq 1)$.

Theorem 6.2. Suppose that $\nu_N \rightarrow \nu$ weakly. The sequence of probability measures $(p_{\nu_N}^N, N \geq 1)$ obeys a large deviations principle in the space $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))^{2r})$, with speed N , and rate function $S_{[0,T]}(\mu|\nu)$ given by (173).

Moreover, if a path $\mu \in (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$ satisfies $S_{[0,T]}(\mu|\nu) < \infty$, then μ_j^c and μ_j^p are absolutely continuous and there exist families of rate matrices $L_{j,c}(t) = \left(l_{z,z'}^{j,c}(t), (z, z') \in \mathcal{E} \right)$ and $L_{j,p}(t) = \left(l_{z,z'}^{j,p}(t), (z, z') \in \mathcal{E} \right)$ such that, for all $1 \leq j \leq r$ and $t \in [0, T]$,

$$\begin{aligned} \dot{\mu}_j^c(t) &= L_{j,c}(t)^* \mu_j^c(t), \\ \dot{\mu}_j^p(t) &= L_{j,p}(t)^* \mu_j^p(t). \end{aligned}$$

Furthermore, in this case, the good rate function $S_{[0,T]}(\mu|\nu)$ is given by

$$\begin{aligned} & \sum_{j=1}^r \left[\alpha_j p_j^c \int_0^T \left(\sum_{(z,z') \in \mathcal{E}} \left(\mu_j^c(t)(z) \lambda_{j,z,z'}^c \left(\mu_j^c(t), \mu_j^p(t) \right) \tau^* \left(\frac{l_{z,z'}^{j,c}(t)}{\lambda_{j,z,z'}^c \left(\mu_j^c(t), \mu_j^p(t) \right)} - 1 \right) \right) \right) dt \right. \\ & \quad \left. + \alpha_j p_j^p \int_0^T \left(\sum_{(z,z') \in \mathcal{E}} \left(\mu_j^p(t)(z) \lambda_{j,z,z'}^p \left(\mu_j^c(t), \mu_1^p(t), \dots, \mu_r^p(t) \right) \right. \right. \right. \\ & \quad \left. \left. \left. \tau^* \left(\frac{l_{z,z'}^{j,p}(t)}{\lambda_{j,z,z'}^p \left(\mu_j^c(t), \mu_1^p(t), \dots, \mu_r^p(t) \right)} - 1 \right) \right) \right) dt \right]. \end{aligned} \tag{174}$$

Proof. We first use a contraction argument to derive a large deviations principle for the sequence $(p_{\nu_N}^N, N \geq 1)$. From Theorem 6.1, the sequence $(P_{\nu_N}^N, N \geq 1)$ obeys a large deviations principle with speed N and rate function $I(\mathbf{Q})$ given by

$$I(\mathbf{Q}) = \begin{cases} \sum_{j=1}^r \left[\alpha_j p_j^c H \left(Q_j^c | R^c \left(\pi(Q_j^c), \pi(Q_j^p) \right) \right) \right. & \text{if } \mathbf{Q} \circ \pi_0^{-1} = \nu, \\ \quad \left. + \alpha_j p_j^p H \left(Q_j^p | R^p \left(\pi(Q_j^c), \pi(Q_1^p), \dots, \pi(Q_r^p) \right) \right) \right] & \\ +\infty & \text{otherwise.} \end{cases}$$

Moreover, from Lemma 6.4, the projection

$$\pi : (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r} \rightarrow (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$$

is continuous at each $\mathbf{Q} \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}$ where $L(\mathbf{Q}) < \infty$, and thus at any \mathbf{Q} such that $I(\mathbf{Q}) < \infty$. The latter corresponds to the effective domain $\mathcal{D}_I = \{\mathbf{Q} : I(\mathbf{Q}) < \infty\}$ of the rate function I (see [29, p. 4]). Therefore, by applying the contraction principle to the large deviations principle of $(P_{\nu_N}^N, N \geq 1)$ (see [29, Theorem 4.2.1, Remark (c)]) with rate I , we deduce that the family of probability measures $(P_{\nu_N}^N \circ \pi^{-1}, N \geq 1)$ obeys a large deviations principle in $(\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$ with the rate function defined, for any $\mu \in (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$, by

$$V(\mu) := \inf \left\{ I(\mathbf{Q}), \mathbf{Q} \in (\mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})))^{2r}, \pi(\mathbf{Q}) = \mu \right\}. \tag{175}$$

We now derive another representation for the rate function V following [24, 49]. Fix $\mu = (\mu_j^c, \mu_j^p, 1 \leq j \leq r) \in (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$. Note that writing $\pi(\mathbf{Q}) = \mu$, with $\mathbf{Q} \in$

$(\mathcal{M}_1(D([0, T], \mathcal{Z})))^{2r}$, is equivalent to $\pi(Q_j^c) = \mu_j^c$ and $\pi(Q_j^p) = \mu_j^p$ for all $1 \leq j \leq r$. Therefore V can be rewritten as

$$V(\mu) = \inf \left\{ \sum_{j=1}^r \left[\alpha_j p_j^c H(Q_j^c | R^c(\mu_j^c, \mu_j^p)) + \alpha_j p_j^p H(Q_j^p | R^p(\mu_j^c, \mu_1^p, \dots, \mu_r^p)) \right], \right. \\ \left. \mathbf{Q} \in (\mathcal{M}_1(D([0, T], \mathcal{Z})))^{2r}, \pi(\mathbf{Q}) = \mu \right\}.$$

Fix $1 \leq j \leq r$. Let $(X_{j,c}^{(i)})_{i \geq 1}$ and $(X_{j,p}^{(i)})_{i \geq 1}$ be sequences of i.i.d. processes with common laws $R^c(\mu_j^c, \mu_j^p)$ and $R^p(\mu_j^c, \mu_1^p, \dots, \mu_r^p)$, respectively. By Sanov’s theorem, the empirical measures

$$\frac{1}{N_j^c} \sum_{i=1}^{N_j^c} X_{j,c}^{(i)} \quad \text{and} \quad \frac{1}{N_j^p} \sum_{i=1}^{N_j^p} X_{j,p}^{(i)}$$

obey large deviations principles as $N_j^c \rightarrow \infty$ and $N_j^p \rightarrow \infty$, with speeds N_j^c and N_j^p , respectively, and rate functions given by

$$Q \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})) \rightarrow H(Q | R^c(\mu_j^c, \mu_j^p))$$

and

$$Q \in \mathcal{M}_1(\mathcal{D}([0, T], \mathcal{Z})) \rightarrow H(Q | R^p(\mu_j^c, \mu_1^p, \dots, \mu_r^p)),$$

respectively. Using the same arguments as in the proof of Lemma 6.4, one can show that the projection π is continuous at any $\mathbf{Q} \in (\mathcal{M}_1(D([0, T], \mathcal{Z})))^{2r}$ such that

$$\sum_{j=1}^r \left[\alpha_j p_j^c H(Q_j^c | R^c(\mu_j^c, \mu_j^p)) + \alpha_j p_j^p H(Q_j^p | R^p(\mu_j^c, \mu_1^p, \dots, \mu_r^p)) \right] < \infty.$$

Thus, the component projections $\pi(Q_j^c)$ and $\pi(Q_j^p)$ are also continuous. Hence, using the contraction principle ([29, Theorem 4.2.1]), the sequences

$$\left\{ t \in [0, T] \rightarrow \frac{1}{N_j^c} \sum_{i=1}^{N_j^c} X_{j,c}^{(i)}(t); N_j^c \geq 1 \right\}$$

and

$$\left\{ t \in [0, T] \rightarrow \frac{1}{N_j^p} \sum_{i=1}^{N_j^p} X_{j,p}^{(i)}(t); N_j^p \geq 1 \right\}$$

obey large deviations principles with speeds N_j^c and N_j^p , respectively, and rate functions

$$\eta \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})) \rightarrow \mathcal{S}_\mu^{j,c}(\eta) := \inf \left\{ H(Q | R^c(\mu_j^c, \mu_j^p)), \right. \\ \left. Q \in \mathcal{M}_1(D([0, T], \mathcal{Z})), \pi(Q) = \eta \right\}$$

and

$$\eta \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})) \rightarrow S_\mu^{j,p}(\eta) := \inf \left\{ H(Q | R^p(\mu_j^c, \mu_1^p, \dots, \mu_r^p)), \right. \\ \left. Q \in \mathcal{M}_1(D([0, T], \mathcal{Z})), \pi(Q) = \eta \right\},$$

respectively. Note that, by using an independence argument and following the same steps as in the proof of Lemma 6.1, one can show that the sequence

$$\left(t \in [0, T] \rightarrow \left(\frac{1}{N_1^c} \sum_{i=1}^{N_1^c} X_{1,c}^{(i)}(t), \frac{1}{N_1^p} \sum_{i=1}^{N_1^p} X_{1,p}^{(i)}(t), \dots, \frac{1}{N_r^c} \sum_{i=1}^{N_r^c} X_{r,c}^{(i)}(t), \frac{1}{N_r^p} \sum_{i=1}^{N_r^p} X_{r,p}^{(i)}(t) \right) \right)_{N \geq 1}$$

obeys a large deviations principle with speed N and rate function

$$\eta = (\eta_1^c, \eta_1^p, \dots, \eta_r^c, \eta_r^p) \in (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r} \rightarrow S_\mu(\eta) = \sum_{j=1}^r \left[\alpha_j p_j^c S_\mu^{j,c}(\eta_j^c) + \alpha_j p_j^p S_\mu^{j,p}(\eta_j^p) \right].$$

In addition, the vector

$$\left(\frac{1}{N_1^c} \sum_{i=1}^{N_1^c} X_{1,c}^{(i)}, \frac{1}{N_1^p} \sum_{i=1}^{N_1^p} X_{1,p}^{(i)}, \dots, \frac{1}{N_r^c} \sum_{i=1}^{N_r^c} X_{r,c}^{(i)}, \frac{1}{N_r^p} \sum_{i=1}^{N_r^p} X_{r,p}^{(i)} \right)_{N \geq 1}$$

obeys a large deviations principle with rate $I(Q)$. Therefore, by a contraction argument and using again the continuity of the projection, we find that

$$\left(t \in [0, T] \rightarrow \left(\frac{1}{N_1^c} \sum_{i=1}^{N_1^c} X_{1,c}^{(i)}(t), \frac{1}{N_1^p} \sum_{i=1}^{N_1^p} X_{1,p}^{(i)}(t), \dots, \frac{1}{N_r^c} \sum_{i=1}^{N_r^c} X_{r,c}^{(i)}(t), \frac{1}{N_r^p} \sum_{i=1}^{N_r^p} X_{r,p}^{(i)}(t) \right) \right)_{N \geq 1}$$

obeys a large deviations principle with rate $V(\mu)$. Hence, by the uniqueness of the rate function (cf. [30, Lemma 2.1.1]), we find

$$V(\mu) = S_\mu(\mu). \tag{176}$$

We next derive another representation for $S_\mu(\nu)$. For any $1 \leq j \leq r$ and $\nu \in \mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z}))$, we have from [49, p. 319] that

$$S_\mu^{j,c}(\nu) = U_\mu^{j,c}(\nu_t(dz)dt), \\ S_\mu^{j,p}(\nu) = U_\mu^{j,p}(\nu_t(dz)dt), \tag{177}$$

where, for all $\tilde{\nu} \in \mathcal{M}_1([0, T] \times \mathcal{Z})$, $U_\mu^{j,c}(\tilde{\nu})$ and $U_\mu^{j,p}(\tilde{\nu})$ are given by the following (see [49, Equation (3.14)]):

$$U_\mu^{j,c}(\tilde{\nu}) = \sup_{f \in \mathcal{C}_1^c} \left\{ \int_0^T \left\langle - \left(\frac{\partial}{\partial t} + A_{\mu(t)}^{j,c} \right) f(t, z) \right. \right. \\ \left. \left. - \sum_{z' : (z, z') \in \mathcal{E}} \tau(f(t, z') - f(t, z)) \lambda_{j,z,z'}^c(\mu_j^c(t), \mu_j^p(t)), \nu_t(dz) \right\rangle \right\}, \tag{178}$$

$$U_{\mu}^{j,p}(\tilde{v}) = \sup_{f \in C_1^c} \left\{ \int_0^T \left\langle - \left(\frac{\partial}{\partial t} + A_{\mu(t)}^{j,p} \right) f(t, z) - \sum_{z': (z, z') \in \mathcal{E}} \tau(f(t, z') - f(t, z)) \lambda_{j, z, z'}^p \left(\mu_j^c(t), \mu_1^p(t), \dots, \mu_r^p(t) \right), \nu_t(dz) \right\rangle \right\}, \tag{179}$$

where C_1^c stands for the set of all continuous functions with compact support on $[0, T] \times \mathcal{Z}$ which are t -differentiable. Using (177), (178), and (179) together with [49, Lemma 3.2], we obtain

$$S_{\mu}^{j,c}(\mu) = \int_0^T \left\| \dot{\mu}_j^c(t) - A_{\mu(t)}^{j,c*} \mu_j^c(t) \right\|_{\mu(t)} dt, \tag{180}$$

$$S_{\mu}^{j,p}(\mu) = \int_0^T \left\| \dot{\mu}_j^p(t) - A_{\mu(t)}^{j,p*} \mu_j^p(t) \right\|_{\mu(t)} dt.$$

Finally, using (176), we deduce that $(p_{v_N}^N, N \geq 1)$ obeys a large deviations principle with rate N and good rate function (173). The representation (174) follows immediately from [49, Lemma 3.2], and the statement about absolute continuity follows from [49, Theorem 3.1]. The theorem is proved. \square

The following result shows that the large deviations principle for $(p_v^N, N \geq 1)$ holds uniformly in the initial condition.

Corollary 6.1. *For any compact set $K \subset (\mathcal{M}_1(\mathcal{Z}))^{2r}$, any closed set $F \subset (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$, and any open set $G \subset (\mathcal{D}([0, T], \mathcal{M}_1(\mathcal{Z})))^{2r}$, we have*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \sup_{\nu \in K} p_v^N(\mu_N \in F) \leq - \inf_{\nu \in K} \inf_{\mu \in F} S_{[0,T]}(\mu|\nu), \tag{181}$$

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \log \inf_{\nu \in K} p_v^N(\mu_N \in G) \geq - \sup_{\nu \in K} \inf_{\mu \in G} S_{[0,T]}(\mu|\nu). \tag{182}$$

Proof. This follows immediately from [29, Corollary 5.6.15] and Theorem 6.2. \square

Acknowledgements

We would like to thank the anonymous referees and the associate editor for having read the paper with great care and made several very important comments that improved the exposition.

Funding information

This research was supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grants and by Carleton University.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] AGLIARI, E., MIGLIOZZI, D. AND TANTARI, D. (2018). Non-convex multi-species Hopfield models. *J. Statist. Phys.* **172**, 1247–1269.
- [2] AKHIL, P. T., ALTMAN, E. AND SUNDARESAN, R. (2019). *A mean-field approach for controlling singularly perturbed multi-population SIS epidemics*. Preprint. Available at <https://arxiv.org/abs/1902.05713>.
- [3] ALBERICI, D., CAMILLI, F., CONTUCCI, P. AND MINGIONE, E. (2021). The multi-species mean-field spin-glass on the Nishimori line. *J. Statist. Phys.* **182**, article no. 2.
- [4] ALEANDRI, M. AND MINELLI, I. G. (2019). Opinion dynamics with Lotka–Volterra type interactions. *Electron. J. Probab.* **24**, 31 pp.
- [5] BARRA, A., CONTUCCI, P., MINGIONE, E. AND TANTARI, D. (2015). Multi-species mean field spin glasses. Rigorous results. *Ann. Inst. H. Poincaré Prob. Statist.* **16**, 691–708.
- [6] BAYRAKTAR, E., CHAKRABORTY, S. AND WU, R. (2020). *Graphon mean field systems*. Preprint. Available at <https://arxiv.org/abs/2003.13180>.
- [7] BAYRAKTAR, E. AND WU, R. (2021). Mean field interaction on random graphs with dynamically changing multi-color edges. *Stoch. Process. Appl.* **141**, 197–244.
- [8] BENAÏM, M. AND LE BOUDEC, J.-Y. (2008). A class of mean field interaction models for computer and communication systems. *Performance Evaluation* **65**, 823–838.
- [9] BHAMIDI, S., BUDHIRAJA, A. AND WU, R. (2019). Weakly interacting particle systems on inhomogeneous random graphs. *Stoch. Process. Appl.* **129**, 2174–2206.
- [10] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
- [11] BOLLEY, F. (2008). Separability and completeness for the Wasserstein distance. In *Séminaire de Probabilités XXI*, eds C. Donati-Martin, M. Émery, A. Rouault and C. Stricker, Springer, Berlin, Heidelberg, pp. 371–377.
- [12] BORKAR, V. S. AND SUNDARESAN, R. (2012). Asymptotics of the invariant measure in mean field models with jumps. *Stoch. Systems* **2**, 322–380.
- [13] BOSSY, M., FAUGERAS, O. AND TALAY, D. (2015). Clarification and complement to ‘Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons’. *J. Math. Neurosci.* **5**, article no. 31, 23 pp.
- [14] BUCKDAHN, R., LI, J. AND PENG, S. (2014). Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents. *SIAM J. Control Optimization* **52**, 451–492.
- [15] BUDHIRAJA, A., MUKHERJEE, D. AND WU, R. (2019). Supermarket model on graphs. *Ann. Appl. Probab.* **29**, 1740–1777.
- [16] BUDHIRAJA, A. AND WU, R. (2016). Some fluctuation results for weakly interacting multi-type particle systems. *Stoch. Process. Appl.* **126**, 2253–2296.
- [17] CARMONA, R. AND ZHU, X. (2016). A probabilistic approach to mean field games with major and minor players. *Ann. Appl. Probab.* **26**, 1535–1580.
- [18] CHONG, C. AND KLÜPPELBERG, C. (2019). Partial mean field limits in heterogeneous networks. *Stoch. Process. Appl.* **129**, 4998–5036.
- [19] COLLET, F. (2014). Macroscopic limit of a bipartite Curie–Weiss model: a dynamical approach. *J. Statist. Phys.* **157**, 1309–1319.
- [20] COLLET, F., FORMENTIN, M. AND TOVAZZI, D. (2016). Rhythmic behavior in a two-population mean-field Ising model. *Phys. Rev. E* **94**, article no. 042139.
- [21] DAWSON, D. A. (1983). Critical dynamics and fluctuations for a mean field model of cooperative behaviour. *J. Statist. Phys.* **41**, 29–85.
- [22] DAWSON, D. A. (1991). Measure-valued Markov processes. In *École d’Été de Probabilités de Saint-Flour XXI—1991*, Springer, Berlin, Heidelberg, pp. 1–260.
- [23] DAWSON, D. A. (2017). *Introductory lectures on stochastic population systems*. Preprint. Available at <https://arxiv.org/abs/1705.03781>.
- [24] DAWSON, D. A. AND GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20**, 247–308.
- [25] DAWSON, D. A., SID-ALI, A. AND ZHAO, Y. Q. (2022). Large-time behavior of finite-state mean-field systems with multi-classes. *Stoch. Systems* **13**, 93–127.
- [26] DAWSON, D. A., TANG, J. AND ZHAO, Y. Q. (2005). Balancing queues by mean field interaction. *Queueing Systems* **49**, 335–361.
- [27] DAWSON, D. A. AND ZHENG, X. (1991). Law of large numbers and central limit theorem for unbounded jump mean-field models. *Adv. Appl. Math.* **12**, 293–326.
- [28] DELATTRE, S., GIACOMIN, G. AND LUÇON, E. (2016). A note on dynamical models on random graphs and Fokker–Planck equations. *J. Statist. Phys.* **165**, 785–798.
- [29] DEMBO, A. AND ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications*, 2nd edn. Springer, Berlin, Heidelberg.

- [30] (1989). Deuschel, J.-D. and Stroock, D. W. *Large Deviations*. Academic Press, New York.
- [31] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.
- [32] FENG, S. (1994). Large deviations for empirical process of mean-field interacting particle system with unbounded jumps. *Ann. Prob.* **22**, 2122–2151.
- [33] FENG, S. (1994). Large deviations for Markov processes with mean field interaction and unbounded jumps. *Prob. Theory Relat. Fields* **100**, 227–252.
- [34] FINNOFF, W. (1993). Law of large numbers for a general system of stochastic differential equations with global interaction. *Stoch. Process. Appl.* **46**, 153–182.
- [35] FINNOFF, W. (1994). Law of large numbers for a heterogeneous system of stochastic differential equations with strong local interaction and economic applications. *Ann. Appl. Prob.* **4**, 494–528.
- [36] FREIDLIN, M. I. AND WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd edn. Springer, Berlin, Heidelberg.
- [37] GÄRTNER, J. (1988). On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137**, 197–248.
- [38] GIESECKE, K., SPILIOPOULOS, K., SOWERS, R. B. AND SIRIGNANO, J. A. (2015). Large portfolio asymptotics for loss from default. *Math. Finance* **25**, 77–114.
- [39] GRAHAM, C. (2000). Chaoticity on path space for a queueing network with selection of the shortest queue amongst several. *J. Appl. Prob.* **37**, 198–211.
- [40] GRAHAM, C. (2008). Chaoticity for multiclass systems and exchangeability within classes. *J. Appl. Prob.* **45**, 1196–1203.
- [41] GRAHAM, C. AND MÉLÉARD, S. (1993). Propagation of chaos for a fully connected loss network with alternate routing. *Stoch. Process. Appl.* **44**, 159–180.
- [42] GRAHAM, C. AND MÉLÉARD, S. (1995). Dynamic asymptotic results for a generalized star-shaped loss network. *Ann. Appl. Prob.* **5**, 666–680.
- [43] GRAHAM, C. AND ROBERT, P. (2009). Interacting multi-class transmissions in large stochastic networks. *Ann. Appl. Prob.* **19**, 2334–2361.
- [44] HWANG, C.-R. AND SHEU, S.-J. (1990). Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker–Planck operators and simulated annealing. *Acta Appl. Math.* **19**, 253–295.
- [45] KAC, M. (1956). Foundations of kinetic theory. In *Proc. 3rd Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Contributions to Astronomy and Physics*, University of California Press, Berkeley, pp. 171–197.
- [46] KIRSCH, W. AND TOTH, G. (2020). Two groups in a Curie–Weiss model with heterogeneous coupling. *J. Theoret. Prob.* **33**, 2001–2026.
- [47] KLEY, O., KLÜPPELBERG, C. AND REICHEL, L. (2015). Systemic risk through contagion in a core-periphery structured banking network. *Banach Center Publ.* **104**, 133–149.
- [48] KNÖPFEL, H., LÖWE, M., SCHUBERT, K. AND SINULIS, A. (2020). Fluctuation results for general block spin Ising models. *J. Statist. Phys.* **178**, 1175–1200.
- [49] LÉONARD, C. (1995). Large deviations for long range interacting particle systems with jumps. *Ann. Inst. H. Poincaré Prob. Statist.* **31**, 289–323.
- [50] LÖWE, M. AND SCHUBERT, K. (2018). Fluctuations for block spin Ising models. *Electron. Commun. Prob.* **23**, 12 pp.
- [51] MCKEAN, H. P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. USA* **56**, 1907–1911.
- [52] MCKEAN, H. P. (1966). Speed of approach to equilibrium for Kac’s caricature of a Maxwellian gas. *Arch. Rational Mech. Anal.* **21**, 343–367.
- [53] MÉLÉARD, S. AND BANSAYE, V. (2015). *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. Springer, Cham.
- [54] MEYLAHN, J. M. (2020). Two-community noisy Kuramoto model. *Nonlinearity* **33**, 1847–1880.
- [55] MITZENMATHER, M. (1996). The power of two choices in randomized load balancing. Doctoral Thesis, University of California, Berkeley.
- [56] NAGASAWA, M. AND TANAKA, H. (1987). Diffusion with interactions and collisions between coloured particles and the propagation of chaos. *Prob. Theory Relat. Fields* **74**, 161–198.
- [57] NAGASAWA, M. AND TANAKA, H. (1987). On the propagation of chaos for diffusion processes with drift coefficients not of average form. *Tokyo J. Math.* **10**, 403–418.
- [58] NGUYEN, D. T., NGUYEN, S. L. AND DU, N. H. (2020). On mean field systems with multiclass. *Discrete Continuous Dynam. Systems* **40**, 683–707.
- [59] PORTER, M. A., ONNELA, J. P. AND MUCHA, P. J. (2009). Communities in networks. *Notices Amer. Math. Soc.* **56**, 1082–1097.

- [60] SKOROKHOD, A. V. (1989). *Asymptotic Methods in the Theory of Stochastic Differential Equations*. American Mathematical Society, Providence, RI.
- [61] SZNITMAN, A. S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, ed. P. L. Hennequin, Springer, Berlin, Heidelberg, pp. 165–251.
- [62] TOUBOUL, J. (2014). Propagation of chaos in neural fields. *Ann. Appl. Prob.* **24**, 1298–1328.
- [63] TOUBOUL, J. (2018). Erratum: 'Propagation of chaos in neural fields'. *Ann. Appl. Prob.* **28**, 3287–3289.
- [64] VARADHAN, S. R. S. (1996). Asymptotic probabilities and differential equations. *Commun. Pure Appl. Math.* **19**, 261–286.
- [65] VVEDENSKAYA, N. D., DOBRUSHIN, R. L. AND KARPELEVICH, F. I. (1996). Queueing system with selection of the shortest of two queues: an asymptotic approach. *Problems Inf. Transmission* **32**, 15–27.
- [66] VVEDENSKAYA, N. D. AND SUHOV, Y. M. (1997). Dobrushin's mean-field approximation for a queue with dynamic routing. *Markov Process. Relat. Fields* **3**, 493–526.
- [67] YASODHARAN, S. AND SUNDARESAN, R. (2022). Large-time behaviour and the second eigenvalue problem for finite-state mean-field interacting particle systems. *Adv. Appl. Prob.* **55**, 85–125.