

CONDITIONAL PROPAGATION OF CHAOS IN A SPATIAL STOCHASTIC EPIDEMIC MODEL WITH COMMON NOISE

MAXIME HAURAY, ÉTIENNE PARDOUX AND YEN V. VUONG

ABSTRACT. We study a stochastic spatial epidemic model where the N individuals carry two features: a position and an infection state, interact and move in \mathbb{R}^d . In this Markovian model, the evolution of the infection states are described with the help of the Poisson Point Processes, whereas the displacement of the individuals are driven by mean field advection, a (state dependence) diffusion and also a common noise, so that the spatial dynamic is a random process. We prove that when the number N of individual goes to infinity, the conditional propagation of chaos holds: conditionally to the common noise, the individuals are asymptotically independent and the stochastic dynamic converges to a "random" nonlinear McKean-Vlasov process. As a consequence, the associated empirical measure converges to a measure, which is solution of a stochastic mean-field PDE driven by the common noise.

Key Words. Stochastic epidemic model, spatial epidemic model, conditional propagation of chaos, mean field limit.

1. INTRODUCTION

Epidemic models have been studied for a long while, in both deterministic and stochastic settings. In this paper, we study a spatial model based on the famous SIR model, the letter S, I and R standing for the different states of an individual which can pass from the compartment of "Susceptible" to the "Infected" one and eventually to the compartment of "Recovered" when the individual recovers from the illness. In our spatial model, an individual will be characterized by:

- Its state $E \in \{S, I, R\} = \{0, 1, 2\}$, since we will identify S with 0, I with 1 and R with 2 in order to simplify the mathematical description,
- Its position, a continuous variable $X \in \mathbb{R}^d$. Typically, $d = 1, 2$ for the propagation of epidemics¹.

So the individual phase space is $\Pi := \mathbb{R}^d \times \{0, 1, 2\}$. We consider a community of N individuals, denoting by $Z_t^{i,N} = (X_t^{i,N}, E_t^{i,N})$, $i = 1, \dots, N$, the position and state of the i^{th} individual at time $t > 0$. The full vector $(Z_t^{i,N})_{i \leq N}$ belongs to the full phase space $\Pi^N = \mathbb{R}^{dN} \times \{0, 1, 2\}^N$.

The introduction of spatial variables will complicate the standard homogeneous SIR model in two directions: by using an infection rate that depends on the position of the individuals, and by taking into account the individual displacements.

An infection rate depending on the position. It is quite clear that in realistic situation, an infected individual will infect a close neighbor with a higher rate than an individual living far away. While these different behaviors are averaged in a homogeneous SIR model, in spatial models we use an infection rate depending on the positions. The infection rate between locations x and $y \in \mathbb{R}^d$ will be given by a function $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_+$, that we will assume to be bounded

¹but the case $d = 3$ is also considered in quite similar models in chemistry for the microscopical description of chemical reactions, see [19].

and Lipschitz. Averaging over all the infected individuals, the susceptible individual i becomes infected (in other words its state jumps from 0 to 1) at time t at the rate

$$\frac{1}{N} \sum_{j=1}^N K(X_t^{i,N}, X_t^{j,N}) \mathbb{1}_{\{E_t^{i,N}=S\}} \mathbb{1}_{\{E_t^{j,N}=I\}}. \quad (1.1)$$

What could be the form of a realistic function K ? For several sedentary species, it could be reasonable to think that the individuals could only infect their close neighbors, and use a compactly supported and constant kernel, of the form

$$K(x, y) = \frac{\beta}{c_d r_0^d} \mathbb{1}_{|x-y| \leq r_0},$$

where r_0 is the size of the "territory" of the individuals, and c_d is the volume of the unit ball, and β is the infection rate.

For the propagation of an epidemic among a human population, it is possible to incorporate in K effects due to the human mobility, which has become a subject of intensive research in the past decades. Most popular models for the human mobility are the so-called gravity model [10] which state that the total number of travel between two cities is inversely proportional to the inverse of the square of the distance between them², and the radiation model [23], which state that the same number depends also on the distribution of population between the two locations. The gravity model naturally leads in our case in the use of a function K of the form

$$K(x, y) = \frac{c}{d^2 + |x - y|^2},$$

with $c, d \in \mathbb{R}_+$ two constants to be fitted with the observations. According to the study [21], that choice is in a good accordance with the observed data on the spread of the Covid epidemic in France in 2020. We refer to [2] for the description of other mobility functions and the study of their relevance in (discrete) spatial epidemic model. To be complete, we also mention a recent study [22] that deduces from the observations a quite general law in urban mobility and that could lead to a general kernel, not only depending on the relative distance between x and y , but also on the attractivity of the different locations. The choice of an adapted kernel is clearly a relevant question, which is still a subject under investigation, and here we will consider a general kernel, depending on the two positions.

Recovery rate. The infectious individuals recover (in other words their state jumps from 1 to 2) at rate $\gamma > 0$ independently of everything else (of the other individuals, the number of infected and the respective positions of the individuals,...) and once an individual recovers, it can not become infected anymore.

Movements of individuals. In order to study the propagation of epidemics among non sedentary species, we will take into account some kind of displacements of the individuals. In our model, each individual moves in \mathbb{R}^d according to

- A mean field interaction (with all the others individuals),
- Independent diffusions,
- A random drift (or diffusion), common to all the individuals.

So the evolution of the position $X_t^{i,N}$ of the i^{th} individual satisfies the following equation

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N V(X_t^{i,N}, E_t^{i,N}, X_t^{j,N}, E_t^{j,N}) dt + \sigma(X_t^{i,N}, E_t^{i,N}) dB_t^i + \sigma_0(X_t^{i,N}, E_t^{i,N}) dB_t^0, \quad (1.2)$$

²Some versions allow to fit the power of the distance rather than using the power 2, or even to replace it by a more general function of the distance.

where the interaction kernel V , the diffusion strength σ and the strength σ_0 of the random drift-diffusion all depend on the individual's state and position, and are bounded Lipschitz continuous with respect to the position variables. Of course, this equation has a meaning on a probability space endowed with the requested Brownian motions (and Poisson Point Processes for the jump-infectious part of the dynamic, as will be introduced below). The Lipschitz hypothesis will be very useful to build a correct theory of existence, uniqueness to that system, and also for our results concerning the large population limit (i.e. when N goes to infinity).

Before going on, let us discuss a bit the possible application of that kind of dynamics, and the particular choice of diffusion, mean-field and common drift-diffusion that could be relevant.

Mean-field interactions are commonly chosen in order to describe swarming or flocking behavior of some species, see for instance [6]. Such phenomena occur on time scales shorter than the one of epidemic propagation. We choose to use it in order to illustrate that our model is compatible with the introduction of a kind of collective interaction.

The common drift term is here written in a quite simple form. It could probably be generalized to a more general random drift without diffusion which is probably the most interesting case) like in [9] of the form

$$\sum_{k=1}^{\infty} \sigma_k(X_t^{i,N}, E_t^{i,N}) dB_t^{0,k},$$

with divergence-free σ_k , which is a quite general way to describe homogeneous random drift³.

The individual diffusion term is quite common to model the displacement of individuals. Modeling it with the help of a standard diffusion is a first step, that is probably not very realistic, since individuals usually do not have a Markovian behavior. One more realistic model among human population is for instance the EPR model (Exploration and Preferential Return) introduced in [24].

We stress out the fact that the diffusion strength depends on the infectious state of the individuals is quite interesting for the application. We could think of the propagation of a rabies epidemic, where the disease affects the displacement behavior of the individuals. But also in the case of epidemics in human population, as the individuals are exchangeable, we may study the case where the susceptible population is at rest, while the infectious individuals follow a diffusion in order to model the dispersion of the epidemic. In that case, the classical diffusion could be an interesting choice, even if some more realistic models should probably include Lévy flights and also be non Markovian, as suggested by the study of banknote diffusion in a population [5].

1.1. The model. In view of the above settings, a description of the epidemiological dynamic with the help of Poisson Point Processes, see Definition 6.1, is suitable. If we choose a probability space equipped with N independent Poisson Point Processes $(P^i)_{i=1,\dots,N}$ and $N + 1$ Brownian motions, the position and state of the individuals will evolve according to the following dynamics:

$$\begin{cases} dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N V(X_t^{i,N}, E_t^{i,N}, X_t^{j,N}, E_t^{j,N}) dt + \sigma(X_t^{i,N}, E_t^{i,N}) dB_t^i + \sigma_0(X_t^{i,N}, E_t^{i,N}) dB_t^0, \\ E_t^{i,N} = E_0^{i,N} + P^i \left(\int_0^t \left\{ \frac{1}{N} \sum_{j \neq i} K(X_s^{i,N}, X_s^{j,N}) \mathbb{1}_{(E_s^{i,N}, E_s^{j,N})=(0,1)} + \gamma \mathbb{1}_{E_s^{i,N}=1} \right\} ds \right). \end{cases} \tag{1.3}$$

Recall that the processes $E_t^{i,N}$ take values in $\{0, 1, 2\}$, since S has been identified with 0, I with 1 and R with 2.

³It is shown in [9] that under the divergence free hypothesis, the Ito and Stratonovich formulation are equivalent, so that we have the right to call it random drift here.

This model is inspired by the previous work [3] where the case without common random drift and mean-field interaction is treated, and results about the Law of Large Number and the Central Limit Theorem are obtained. Here we will generalize the Law of Large Number result.

Let us also mention some previous works in other related contexts where a common noise affects all individuals : interacting particle systems with a common random drift [9], mean field games [7], [8].

An important concept in studying the mean field limit of the interacting particle system is the propagation of chaos [15, 25], which roughly states that when the number of particles tends to infinity, the chaotic character of the initial distribution propagates through the system at any time. In the present model, the individuals always keep track of the common noise B_t^0 so it is impossible to expect the asymptotically independence at the limit. Nevertheless, it is reasonable to expect that the individuals become asymptotically independent conditionally on the information generated by the common noise. The objective of this paper is to establish the conditional propagation of chaos of the N individuals process towards a random nonlinear (or McKean-Vlasov) process defined later. This also implies the convergence of the empirical measure process

$$t \mapsto \mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{(X_t^{i,N}, E_t^{i,N})},$$

which is a random process with value in $\mathcal{P}(\Pi) := \mathcal{P}(\mathbb{R}^d \times \{0, 1, 2\})$, the set of probabilities on Π . The proof will follow a standard coupling method introduced by McKean in [20].

1.2. The existing literature and the novelty of our result. The idea of our paper was inspired originally mostly by the two papers [3] and [9]. Here we detail the novelties and improvements in our work with respect to these two references.

Our first inspiring paper [3] was written by A. Emakoua, S. Bowong and the third author. They studied a spatial epidemic model which is similar to ours but a bit simpler, without

- the common stochastic case (with our notation $\sigma_0 = 0$),
- interaction among the individuals (with our notation $V = 0$).

In their model, the positions of the individuals are independent Brownian motions. In that quite simpler setting, their main result is a Central Limit Theorem (CLT) in the limit of a large number of individuals. The proof of the CLT result is delicate and technical, and relies of course on a Law of Large Number (LLN) result, much simpler to prove. In our present article we focus on the LLN result, and improve it in two directions:

- We prove it in a more general (and realistic) situation,
- We obtain *quantitative* estimates: in Wasserstein distance rather than a qualitative result stating the weak convergence.

Remark that the stronger CLT result of [3] also gives as a byproduct of the order of convergence in the LLN, but not for the Wasserstein distance as we obtain. We also mention that the proof of a CLT in our setting will be the object of a future work.

Our second inspiring article was written by M. Coghi and F. Flandoli [9]. In that work, the authors prove a conditional propagation of chaos result for an interacting particle system where the particles (having only a position parameter and no distinct types) are subject to individual and common noises. With our formalism, there are two main differences:

- There is no internal state E : so the dynamical equation in their model is very similar to the first equation in (1.3), used with functions V, σ, σ_0 independent of E .

- They used a quite general (smooth enough) cylindrical Brownian motion with quite general (smooth) correlation function. That noise could be written as a infinite series of the form $dW(s, x) = \sum_k \sigma^k(x) dB_s^{0,k}$, where $(B_s^{0,k})_{s \geq 0}$ are standard Brownian motions and $(\sigma^k)_{k \geq 0}$ a family of smooth enough functions. Their common noises are “richer” and more realistic than ours since we use only *one* standard Brownian motion to define ours. But it seems reasonable to expect that our result extend to the more realistic situation studied in [9].

If we restrict our result to the case where the movement does not depend upon the internal state $(V, \sigma, \sigma_0$ independent of E) so that we can study the dynamics of the position only, and if we use only our simpler type of common noise in [9], then both results are very similar. We both provide the same quantitative estimate of conditional propagation of chaos in Wasserstein distance of order one. The only difference is in the constants appearing in the estimates. But even if the results are similar, there are still two differences that we would like to emphasize.

The first difference is that the use of Wasserstein distance of order one is motivated in [9] by the willingness to use the standard tool in the field, in order to compare with previous results. But in their setting, the use of the Wasserstein distance of order two, which is defined with a quadratic cost and thus behave much better than the order one distance with respect to the diffusion, would have simplified the proof. In our model, the use of the order one Wasserstein distance is mandatory in order to treat the Poissonian part of the epidemiological dynamics. See the Remark 3.2 for a more detailed explanation of that point.

Our proof contains a simplification that could be used also in the work by M. Coghi and F. Flandoli. In order to control the evolution of the Wasserstein distance of order one under diffusion, they used a short time estimate that allows them to iterate in order to get the desired estimate on arbitrary finite time, the iteration step being a bit technical. We perform a more careful analysis of the evolution of the same quantity and get directly an estimate valid for an arbitrary finite time, without any iteration step. See Remark 3.1 in Section 3 for a more precise description of the improvement.

Organisation of the paper. We present the model state our main results in Section 2. The next sections are devoted to the proof of the main results. Section 3 is devoted to the proof of Theorem 2.5 and to a result on the propagation of moments. Section 4 is devoted to the proof of Theorem 2.6 and some discussions on the uniqueness of solutions to the mean field limit equation. Section 5 contains the proof of Theorem 2.7 and Corollary 2.8.

2. PRELIMINARIES AND MAIN RESULTS

2.1. Notations.

Basic notations.

- Variables denoted by x or y (also with capital letters) denotes a spatial position in \mathbb{R}^d ,
- Variables denoted by e (resp. E) or f (resp. F) denotes an epidemiological status in $\{0, 1, 2\}$,
- Variables denoted by a z (resp. Z) denotes a pair (x, e) of a position and an epidemiological status.

The underlying space. Throughout this paper, we keep assuming that the independent noises $(B^i)_{i=1, \dots, N}$ and $(P^i)_{i=1, \dots, N}$ are constructed on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and the common noise B^0 is constructed on another space $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$. We now define the product structure $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, where $\Omega = \tilde{\Omega} \times \Omega^0$, $(\mathcal{F}, \mathbb{P})$ is the completion of $(\tilde{\mathcal{F}} \otimes \mathcal{F}^0, \tilde{\mathbb{P}} \otimes \mathbb{P}^0)$, and $(\mathcal{F}_t)_{t \geq 0}$ is the complete right continuous version of $(\tilde{\mathcal{F}}_t \otimes \mathcal{F}_t^0)_{t \geq 0}$. With this construction, if we consider

Z a Π -valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ then the conditional law of Z given \mathcal{F}^0 , denoted by $\mathcal{L}(Z|\mathcal{F}^0)$, is a $\mathcal{P}(\Pi)$ -valued random variable defined on $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$.

A more compact formulation of the system (1.3). With the convenient notations

$$V_s^{i,N} := \frac{1}{N} \sum_{j \neq i} V(X_s^{i,N}, E_s^{i,N}, X_s^{j,N}, E_s^{j,N}), \quad K_s^{i,N} := \frac{1}{N} \sum_{j \neq i} K(X_s^{i,N}, X_s^{j,N}) \mathbb{1}_{E_{s^-}^{j,N}=1},$$

the system (1.3) becomes

$$\begin{cases} dX_t^{i,N} &= V_t^{i,N} dt + \sigma(X_t^{i,N}, E_t^{i,N}) dB_t^i + \sigma_0(X_t^{i,N}, E_t^{i,N}) dB_t^0, \\ E_t^{i,N} &= E_0^{i,N} + P^i \left(\int_0^t \left\{ K_s^{i,N} \mathbb{1}_{E_{s^-}^{i,N}=0} + \gamma \mathbb{1}_{E_{s^-}^{i,N}=1} \right\} ds \right). \end{cases} \quad (2.1)$$

Poisson Random Measure rather than Poisson Point Process. The counting processes describing the evolution of the type of each individual used in (2.1) can be rewritten with the help of independent standard Poisson Random Measures $(\mathcal{Q}^i)_{i=1,\dots,N}$ on $\mathbb{R}_+ \times \mathbb{R}_+$ (see Definition 6.2 in the Appendix):

$$E_t^{i,N} = E_0^{i,N} + \int_{[0,t] \times \mathbb{R}_+} \mathbb{1}_{\{u \leq K_s^{i,N} \mathbb{1}_0(E_{s^-}^{i,N}) + \gamma \mathbb{1}_1(E_{s^-}^{i,N})\}} \mathcal{Q}^i(ds, du),$$

In fact the integral quantity in the r.h.s. has the same law (trajectorially) as the stochastic term in the r.h.s in the second line of (2.1). So the system of equations (2.1) can be rewritten as follows

$$\begin{cases} dX_t^{i,N} &= V_s^{i,N} dt + \sigma(X_t^{i,N}, E_t^{i,N}) dB_t^i + \sigma_0(X_t^{i,N}, E_t^{i,N}) dB_t^0, \\ E_t^{i,N} &= E_0^{i,N} + \int_{[0,t] \times \mathbb{R}_+} \mathbb{1}_{\{u \leq K_s^{i,N} \mathbb{1}_0(E_{s^-}^{i,N}) + \gamma \mathbb{1}_1(E_{s^-}^{i,N})\}} \mathcal{Q}^i(ds, du). \end{cases} \quad (2.2)$$

Even if these Poisson Random Measures $(\mathcal{Q}^i)_{i=1,\dots,N}$ may seem less simple to handle than the Poisson Point Processes $(P^i)_{i=1,\dots,N}$, it will be of a great help in the sequel when we couple processes. Thanks to these Poisson Random Measures, we will use a true parallel coupling, in the sense that we will construct coupled processes that jump very often exactly at the same time. It is not possible to construct the same coupling by using Poisson Point Processes. In the latter case, our coupled processes would jump at close times but not exactly at the same times. This has an important consequence: we will be able to obtain estimates of sup norms without allowing for small change in time as it is usual in the Skorohod topology. This simplifies a lot the proofs and allow us to give quantitative convergence results.

Proposition 2.1. *Strong existence and uniqueness holds for both system (1.3) and system (2.2). More precisely, given a probability space endowed with the appropriate Brownian motions and Poisson Point Processes in the case of system (1.3) (or Poisson Random Measures in the case of system (2.2)), there exists a unique solution to the system (1.3) and (2.2).*

Moreover, solutions to (1.3) or (2.2) constructed on possibly different probability spaces have the same law.

The proof of Proposition 2.1 relies on standard argument since all the coefficients appearing in the equation are smooth enough. We refer for instance to [16, Theorem 9.1].

Wasserstein distance. We now introduce the Wasserstein distance together with its variants which are the main tools in the estimates for the convergence of the empirical measures throughout this paper.

For $p \geq 1$, denote by $\mathcal{P}_p(\Pi)$ the set of probability measures μ on Π with finite p -th order moment, i.e. satisfying $\int_{\Pi} |z|^p \mu(dx, de) < \infty$. For μ and ν in $\mathcal{P}_p(\Pi)$, we define the p -Wasserstein distance by

$$W_p^p(\mu, \nu) = \inf \left\{ \int_{\Pi \times \Pi} |z - z'|^p m(dz, dz') : m \in \mathcal{P}_p(\Pi \times \Pi) \text{ with marginals } \mu \text{ and } \nu \right\}.$$

It can also be reformulated using random variables as follows,

$$W_p^p(\mu, \nu) = \inf \left\{ \mathbb{E}|Z - Z'|^p \mid (Z, Z') \text{ r.v. s.t. } \mathcal{L}(Z) = \mu \text{ and } \mathcal{L}(Z') = \nu \right\}.$$

This distance induces the topology of weak convergence of measures together with the convergence of all moments of order up to p . When $p = 1$, the 1-Wasserstein distance is equivalent to the bounded Lipschitz distance

$$W_1(\mu, \nu) = \sup \left\{ \int_{\Pi} \phi(x) \mu(dx) - \int_{\Pi} \phi(x) \nu(dx) : \phi : \Pi \rightarrow \mathbb{R} \text{ with } Lip(\phi) \leq 1 \right\}.$$

Traditionally, to deal with the diffusion terms in the equation, it seems convenient to use the 2-Wasserstein distance. But in this work, we have some compelling reasons to use the 1-Wasserstein distance rather than other order p (with $p > 1$) due to the estimates for the jumps on the epidemiological states.

2.2. Conditional propagation of chaos. In our setting, individuals are always linked through the common noise so there is no way to expect any full independence in the limit. However, we expect the conditional independence, given common noise. That is why we need a conditional version of chaoticity. We will reconsider some definitions and basic properties in the classical theory of mean field analysis of interacting particle systems, then adapt them to a frame with the presence of a common environmental noise. Before stating properly the conditional chaoticity, we need a definition of conditional exchangeability. First, recall that a sequence of random variables is said to be exchangeable if its joint distribution is invariant under any finite permutation.

Definition 2.2. A sequence of N random variables $((Z^{i,N})_{i \leq N})_{N \in \mathbb{N}}$ is said to be exchangeable conditionally on \mathcal{F}^0 if, for any permutation σ of $\{1, \dots, N\}$,

$$\mathbb{P}^0 - a.s. \ , \quad \mathcal{L}\left((Z^{\sigma(i),N})_{i \leq N} \mid \mathcal{F}^0\right) = \mathcal{L}\left((Z^{i,N})_{i \leq N} \mid \mathcal{F}^0\right).$$

Definition 2.3. Let μ be a \mathcal{F}^0 -measurable random probability measure on Π or on $C([0, T], \Pi)$. The sequence $((Z^{i,N})_{i \leq N})_{N \in \mathbb{N}}$ of conditionally exchangeable random variables is called μ -chaotic if it satisfies any of three following equivalent properties :

- (i) $\mathbb{P}^0 - a.s. \ , \ \forall k \in \mathbb{N}, \quad \mathcal{L}((Z^{1,N}, \dots, Z^{k,N}) \mid \mathcal{F}^0) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mu^{\otimes k},$
- (ii) $\mathbb{P}^0 - a.s. \ , \quad \mathcal{L}((Z^{1,N}, Z^{2,N}) \mid \mathcal{F}^0) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mu^{\otimes 2},$
- (iii) $\mathbb{P}^0 - a.s. \ , \quad \mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{Z^{i,N}} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mu.$

The above definition is an adaptation of the classical one from [25] (See also [15] for a quantified version). Apparently, the conditional exchangeability is the key property in that definition. We will rely on this assumption in order to prove the equivalence of the three properties in Definition 2.3. The proof is in the Appendix 6.2.

Under the assumption that initially the individuals are conditionally exchangeable, and thanks to the unique solvability of system (1.3), the conditional exchangeability property persists in time. Indeed, we state that property in the following lemma.

Lemma 2.4. *The conditional exchangeability assumption holds trajectoryally, i.e.*

$$\mathbb{P}^0 - a.s., \forall \sigma \in \Sigma^N, \quad \mathcal{L}\left(\left(Z^{\sigma(i),N}\right)_{i \leq N}\right) = \mathcal{L}\left(\left(Z^{i,N}\right)_{i \leq N}\right).$$

where Σ^N denote the set of permutations of $\{1, \dots, N\}$.

In particular, at any time t , for all $\phi \in C_b(\Pi^N)$,

$$\mathbb{E}\left[\phi\left(Z_t^{1,N}, \dots, Z_t^{N,N}\right) \mid \mathcal{F}_t^0\right] = \mathbb{E}\left[\phi\left(Z_t^{\sigma(1),N}, \dots, Z_t^{\sigma(N),N}\right) \mid \mathcal{F}_t^0\right].$$

The proof of Lemma 2.4 can be found in the Appendix 6.3.

2.3. Main results. Before introducing the exact limit of the stochastic interacting system (1.3), let us present a heuristic derivation of how the limit should look like.

The individuals are driven by independent diffusions, a drift term, jump terms with mean-field rate, and a common diffusion. So they will soon be strongly correlated by the action of the common diffusion. A possible limit towards a mean-field equation is necessarily random, due to that common noise.

We can reasonably assume that the individuals will remain almost conditionally independent along the dynamics. In fact, the influence of one individual on another one in the mean field terms $\frac{1}{N} \sum_j V(Y_t^{i,N}, Y_t^{j,N})$ and $\frac{1}{N} \sum_j K(X_t^{i,N}, X_t^{j,N})$ is of order $\frac{1}{N}$ and we can expect that it will remain small. So we can reasonably expect that in the limit $N \rightarrow \infty$, and conditionally to the common noise, the individual i becomes independent from the other ones, and follows an equation similar to (1.3), with the mean field terms replaced by conditional (w.r.t. the common noise) expectations.

In the sequel, we will show that this formal argument leads at least to the right limit. The main result of this paper show that the sequence $(X^{1,N}, \dots, X^{N,N})_{N \in \mathbb{N}}$ is in fact X -chaotic, where X is a process solution to the following non-linear jumping SDE (or McKean-Vlasov jump diffusion process),

$$\begin{cases} dX_t &= V_{\mu_t}(X_t, E_t)dt + \sigma(X_t, E_t)dB_t + \sigma_0(X_t, E_t)dB_t^0, \\ E_t &= E_0 + P \left(\int_0^t \left(K_{\mu_s}(X_s) \mathbb{1}_{E_{s-}=0} + \gamma \mathbb{1}_{E_{s-}=1} \right) ds \right), \\ \mu_t &= \mathcal{L}((X_t, E_t) \mid \mathcal{F}_t^0), \end{cases} \quad (2.3)$$

where μ_t is the conditional law of (X_t, E_t) given \mathcal{F}_t^0 , and the notations $V_\mu : \Pi \rightarrow \mathbb{R}^d$ and $K_\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ are defined as follows,

$$\begin{aligned} V_\mu(x, e) &= \int_{\Pi} V(x, e, y, f) \mu(dy, df), \\ K_\mu(x) &= \int_{\Pi} \mathbb{1}_1(e) K(x, y) \mu(dy, de). \end{aligned} \quad (2.4)$$

First, we state a theorem about the well-posedness of this stochastic system. It is in fact necessary to have a well-posed limit SDE if we want to state a result of convergence when the number of individuals tends to infinity.

Theorem 2.5. *(i) (Strong existence) Given the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Under the Lipschitz continuous assumptions on the kernels K, V and the coefficients σ, σ_0 , there*

exists a solution $Z_t := (X_t, E_t)$ to the SDE (2.3) associated to any given \mathcal{F}_0 -measurable initial condition $Z_0 = (X_0, E_0)$.

(ii) (Strong stability) If $(X_t^1, E_t^1), (X_t^2, E_t^2)$ are two solutions to (2.3) built on the same probability space with the same driving noises, then for any $t \geq 0$, one has the following stability estimate

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |(X_s^1, E_s^1) - (X_s^2, E_s^2)| \right] \leq C(t) \mathbb{E} \left[|(X_0^1, E_0^1) - (X_0^2, E_0^2)| \right].$$

The derivation of macroscopic equations from microscopic models is a classical topic in the study of interacting particle systems in statistical physics. For this epidemic interacting system, we introduce the Kolmogorov forward equation associated to the nonlinear SDE (2.3), which is the following nonlinear SPDE, where for each $t \geq 0$, μ_t is a random probability on Π :

$$d\mu_t = -\partial_x \cdot (V_{\mu_t} \mu_t) dt + \frac{1}{2} \text{Tr} [\nabla_{xx}^2 ((\sigma \sigma^T + \sigma_0 \sigma_0^T) \mu_t)] dt - \partial_x \cdot (\sigma_0 \mu_t) dB_t^0 + (\Gamma_t^I + \Gamma_t^R) dt, \quad (2.5)$$

where

$$\begin{aligned} \Gamma_t^I(x, e) &= K_{\mu_t}(x) \mu_t(x, 0) (\mathbb{1}_{e=1} - \mathbb{1}_{e=0}), \\ \Gamma_t^R(x, e) &= \gamma \mu_t(x, 1) (\mathbb{1}_{e=2} - \mathbb{1}_{e=1}). \end{aligned}$$

In fact, looking in details at the above equation, we recognize the first term in the r.h.s. as the drift term created by the drift coefficient in (2.3), the second as the diffusion term linked to the individual Brownian motions, the third (with a part of the second one) is the random drift term linked to the common Brownian motion, and the last two Γ terms are related to the jumps between the three different states. We will justify this relation rigorously by the following proposition.

Proposition 2.6. *Assume that X is a solution of equation (2.3), again under the assumption of Theorem 2.5. Then the collection $(\mu_t)_{t \geq 0}$ of its marginal laws solves the above SPDE (2.5).*

In the applications, it is worth to show explicitly the evolution of the empirical measure on each state of the epidemic. We can represent the function $V : (\mathbb{R}^d \times \{0, 1, 2\})^2 \rightarrow \mathbb{R}^d$ as a matrix $(V_{i,j})_{i,j \in \{0,1,2\}}$ where its entries are functions defined on $\mathbb{R}^d \times \mathbb{R}^d$, and introduce the matrix $A[\mu]$ which describes for the epidemiological interactions,

$$A[\mu] := \begin{pmatrix} -K_{\mu} & 0 & 0 \\ K_{\mu} & -\gamma & 0 \\ 0 & \gamma & 0 \end{pmatrix}.$$

If we write μ_t as

$$\mu_t = {}^t(\mu_t \mathbb{1}_{e=0}, \mu_t \mathbb{1}_{e=1}, \mu_t \mathbb{1}_{e=2}) = {}^t(\mu_t^S, \mu_t^I, \mu_t^R),$$

then the above equation (2.5) rewrites

$$d\mu_t = -\partial_x \cdot (V_{\mu_t} \mu_t) dt + \frac{1}{2} \text{Tr} [\nabla_{xx}^2 ((\sigma \sigma^T + \sigma_0 \sigma_0^T) \mu_t)] dt - \partial_x \cdot (\mu_t \sigma_0) dB_t^0 + A[\mu_t] \mu_t dt, \quad (2.6)$$

where the differential operators act on each component of μ_t .

The results about the conditional propagation of chaos. Back to the epidemic dynamic system (1.3), our goal is to prove that this system converges in an appropriate sense to the McKean-Vlasov SDE (2.3). For that purpose, we will need to extend the classical results of propagation of chaos to the conditional one.

Following a standard strategy, we will couple the interacting system to an auxiliary system, made of N i.i.d. copies of the limit SDE, constructed on the same probability space with the

same initial conditions (or at least properly coupled initial conditions) and the same driving random process as for the interacting system.

Then, the estimates used in the proof of the stability result of Theorem 2.5 together with some previously known bound (in Wasserstein sense) on the speed of convergence in the empirical law of large number [12] adapted to the case of conditionally independent random variables in Appendix 6.4 will allow us to establish the following quantitative result about the conditional propagation of chaos for the interacting system of N individuals.

On the same probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ introduced before, for each $i \in \{1, \dots, N\}$, we denote by $\bar{Z}^i = (\bar{X}_t^i, \bar{E}_t^i)$ the strong solution to the McKean-Vlasov jump-diffusion SDE similar to (2.3) but for (independent) driving random processes B^i, P^i instead of B, P and subjects to the initial conditions $\{(\bar{X}_0^i, \bar{E}_0^i)\}_{1 \leq i \leq N}$. Note that by Theorem 2.5, for each i , that SDE has a unique solution and obviously the conditional law $\mathcal{L}((\bar{X}_t^i, \bar{E}_t^i) | \mathcal{F}_t^0)$ is the probability measure μ_t which solves (2.5).

Theorem 2.7. *Assume that the initial conditions $\{(X_0^{i,N}, E_0^{i,N})\}, \{(\bar{X}_0^i, \bar{E}_0^i)\}, i = 1, \dots, N$ are conditionally exchangeable, and have a finite moment of order q for some $q > 2$. Then, there exists a constant $C(t)$ such that, for all $i \in \{1, \dots, N\}$,*

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| \right] \leq C(t) \left(\mathbb{E} \left[|X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i| \right] + \alpha_d(N) \right),$$

where $\alpha_d(N) = N^{-1/2}$ if $d = 1$, $\alpha_d(N) = N^{-1/2} \log(N)$ if $d = 2$, and $\alpha_d(N) = N^{-1/d}$ if $d \geq 3$.

This result is trajectorial. If we only care about the evolution of the empirical measure, then we could get a simpler result expressed in terms of empirical measures only.

Corollary 2.8. *Suppose that the random measure μ_0 has a finite moment of order q for some $q > 2$, and that K, V, σ, σ_0 are Lipschitz continuous (as in Theorem 2.5). Then for all $t > 0$, there exists a constant $C(t)$ such that*

$$\mathbb{E} [W_1(\mu_t^N, \mu_t)] \leq C(t) \left(\mathbb{E} [W_1(\mu_0^N, \mu_0)] + \alpha_d(N) \right),$$

where $\alpha_d(N) = N^{-1/2}$ if $d = 1$, $\alpha_d(N) = N^{-1/2} \log(N)$ if $d = 2$, and $\alpha_d(N) = N^{-1/d}$ if $d \geq 3$.

3. PROOF OF THEOREM 2.5 : WELL-POSEDNESS OF THE NONLINEAR SDE

Proof of ii). We first establish a stability estimate for the solutions of the system (2.3) which is the main estimate for the rest of this paper.

Let $(X_t^1, E_t^1), (X_t^2, E_t^2)$ be two solutions to the SDE (2.3) and μ_t^1, μ_t^2 respectively be conditional law of $(X_t^1, E_t^1), (X_t^2, E_t^2)$ given \mathcal{F}_t^0 .

We have

$$\begin{aligned} X_t^1 - X_t^2 &= X_0^1 - X_0^2 + \int_0^t (V_{\mu_s^1}(X_s^1, E_s^1) - V_{\mu_s^2}(X_s^2, E_s^2)) ds \\ &\quad + \int_0^t (\sigma(X_s^1, E_s^1) - \sigma(X_s^2, E_s^2)) dB_s + \int_0^t (\sigma_0(X_s^1, E_s^1) - \sigma_0(X_s^2, E_s^2)) dB_s^0, \\ E_t^1 - E_t^2 &= E_0^1 - E_0^2 + \int_{[0,t] \times \mathbb{R}_+} \mathbb{1}_{\{u \leq K_{\mu_s^1}(X_s^1) \mathbb{1}_0(E_s^1) + \gamma \mathbb{1}_1(E_s^1)\}} \mathcal{Q}(ds, du) \\ &\quad - \int_{[0,t] \times \mathbb{R}_+} \mathbb{1}_{\{u \leq K_{\mu_s^2}(X_s^2) \mathbb{1}_0(E_s^2) + \gamma \mathbb{1}_1(E_s^2)\}} \mathcal{Q}(ds, du), \end{aligned}$$

which leads to

$$\begin{aligned} \sup_{0 \leq s \leq t} |X_s^1 - X_s^2| &\leq |X_0^1 - X_0^2| + \sup_{0 \leq s \leq t} \left| \int_0^s (V_{\mu_r^1}(X_r^1, E_r^1) - V_{\mu_r^2}(X_r^2, E_r^2)) dr \right| \\ &\quad + \sup_{0 \leq s \leq t} \left| \int_0^s (\sigma(X_r^1, E_r^1) - \sigma(X_r^2, E_r^2)) dB_r \right| \\ &\quad + \sup_{0 \leq s \leq t} \left| \int_0^s (\sigma_0(X_r^1, E_r^1) - \sigma_0(X_r^2, E_r^2)) dB_r^0 \right| \end{aligned}$$

and

$$\sup_{0 \leq s \leq t} |E_s^1 - E_s^2| \leq |E_0^1 - E_0^2| + \sup_{0 \leq s \leq t} \int_{[0, s] \times \mathbb{R}_+} |\mathbb{1}_{\{u \leq \alpha_r^1\}} - \mathbb{1}_{\{u \leq \alpha_r^2\}}| \mathcal{Q}(dr, du),$$

where

$$\begin{aligned} \alpha_r^1 &:= (K_{\mu_r^1}(X_r^1) \mathbb{1}_0(E_r^1) + \gamma \mathbb{1}_1(E_r^1)), \\ \alpha_r^2 &:= (K_{\mu_r^2}(X_r^2) \mathbb{1}_0(E_r^2) + \gamma \mathbb{1}_1(E_r^2)). \end{aligned}$$

We now take the expectation on both sides and our goal is to establish locally uniform in time bounds for the displacement and the jump between the pairs $(X_t^1, E_t^1), (X_t^2, E_t^2)$. For any $s \leq t$, we denote by m_s an arbitrary coupling of (μ_s^1, μ_s^2) on $\Pi \times \Pi$.

Step 1. First, we will give a bound for the drift term in the displacement.

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s (V_{\mu_r^1}(X_r^1, E_r^1) - V_{\mu_r^2}(X_r^2, E_r^2)) dr \right| \right] \\ &\leq \mathbb{E} \left[\int_0^t |V_{\mu_s^1}(X_s^1, E_s^1) - V_{\mu_s^2}(X_s^2, E_s^2)| ds \right] \\ &= \mathbb{E} \left[\int_0^t \left| \int_{\Pi \times \Pi} (V(X_s^1, E_s^1, x, e) - V(X_s^2, E_s^2, y, f)) m_s(dx, de, dy, df) \right| ds \right] \\ &\leq L_V \int_0^t \mathbb{E} \left[|X_s^1 - X_s^2| + |E_s^1 - E_s^2| + \int_{\Pi \times \Pi} (|x - y| + |e - f|) m_s(dx, de, dy, df) \right] ds. \end{aligned}$$

To treat the diffusion terms, we apply the Burkholder–Davis–Gundy inequality and use the Lipschitz continuity of the coefficients σ, σ_0

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s (\sigma(X_r^1, E_r^1) - \sigma(X_r^2, E_r^2)) dB_r \right| + \sup_{0 \leq s \leq t} \left| \int_0^s (\sigma_0(X_r^1, E_r^1) - \sigma_0(X_r^2, E_r^2)) dB_r^0 \right| \right] \\ &\leq C_1(L_\sigma + L_{\sigma_0}) \mathbb{E} \left[\left(\int_0^t (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|)^2 ds \right)^{1/2} \right]. \end{aligned}$$

Remark 3.1. *At this point, a typical difficulty arises where there is no way to remove the powers or simplify the terms. One possibility is to take the supremum in time inside the integral and bound it first on a small enough time interval as in [9] or [11]. Once we have an estimate on a small time interval, the proof is done by iterating over the remain intervals. Nevertheless, in this paper we will make a slight modification by controlling the quantities inside the integral in order to obtain directly an estimate on an arbitrary time interval.*

Indeed, one has

$$\begin{aligned}
& C_1(L_\sigma + L_{\sigma_0}) \mathbb{E} \left[\left(\int_0^t (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|)^2 ds \right)^{1/2} \right] \\
& \leq \mathbb{E} \left[C_1(L_\sigma + L_{\sigma_0}) \left(\sup_{0 \leq s \leq t} (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|) \right)^{1/2} \left(\int_0^t (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|) ds \right)^{1/2} \right] \\
& \leq \frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|) \right] + \frac{C_1^2(L_\sigma + L_{\sigma_0})^2}{2} \int_0^t \mathbb{E} [|X_s^1 - X_s^2| + |E_s^1 - E_s^2|] ds.
\end{aligned}$$

Now summing up all the above terms on both sides, we obtain an appropriate estimate for the displacements

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^1 - X_s^2| \right] & \leq \mathbb{E} [|X_0^1 - X_0^2|] + \frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} (|X_s^1 - X_s^2| + |E_s^1 - E_s^2|) \right] \\
& \quad + \left(L_V + \frac{C_1^2(L_\sigma + L_{\sigma_0})^2}{2} \right) \int_0^t \mathbb{E} [|X_s^1 - X_s^2| + |E_s^1 - E_s^2|] ds \quad (3.1) \\
& \quad + L_V \int_0^t \mathbb{E} \left[\int_{\Pi \times \Pi} (|x - y| + |e - f|) m_s(dx, de, dy, df) \right] ds.
\end{aligned}$$

Step 2. For the Poisson part with the jumps on the epidemiological states, it could be easier to deal with the supremum in time. Indeed, notice that

$$|\mathbb{1}_{u \leq \alpha_r^1} - \mathbb{1}_{u \leq \alpha_r^2}| = \mathbb{1}_{u \in [\alpha_r^1 \wedge \alpha_r^2, \alpha_r^1 \vee \alpha_r^2]}$$

and

$$\alpha_r^1 \vee \alpha_r^2 - \alpha_r^1 \wedge \alpha_r^2 = |\alpha_r^1 - \alpha_r^2|.$$

Using the above identities, we obtain the following estimate

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq s \leq t} \int_{[0, s] \times \mathbb{R}_+} |\mathbb{1}_{\{u \leq \alpha_r^1\}} - \mathbb{1}_{\{u \leq \alpha_r^2\}}| \mathcal{Q}(dr, du) \right] \\
& \leq \mathbb{E} \left[\int_{[0, t] \times \mathbb{R}_+} \mathbb{1}_{\{u \in [\alpha_s^1 \wedge \alpha_s^2, \alpha_s^1 \vee \alpha_s^2]\}} \mathcal{Q}(ds, du) \right] \\
& = \mathbb{E} \left[\int_0^t \left| (K_{\mu_r^2}(X_r^2) \mathbb{1}_0(E_r^2) + \gamma \mathbb{1}_1(E_r^2)) - (K_{\mu_r^1}(X_r^1) \mathbb{1}_0(E_r^1) + \gamma \mathbb{1}_1(E_r^1)) \right| dr \right] \\
& \leq \mathbb{E} \left[\int_0^t |K_{\mu_s^1}(X_s^1) \mathbb{1}_0(E_s^1) - K_{\mu_s^2}(X_s^2) \mathbb{1}_0(E_s^2)| ds \right] + \gamma \mathbb{E} \left[\int_0^t |\mathbb{1}_1(E_s^1) - \mathbb{1}_1(E_s^2)| ds \right] \\
& = I_1 + I_2.
\end{aligned}$$

Remark 3.2. In the first inequality, we can see that no matter the initial degree of the integrand is, we always end with $\mathbb{1}_{u \in [\alpha_r^1 \wedge \alpha_r^2, \alpha_r^1 \vee \alpha_r^2]}$. This is usual when dealing with the Poisson processes. Hence, the following estimates will keep that of order 1 and it is impossible to recover the initial order to get an inequality in type of the Gronwall lemma in the end. Therefore, this crucial point forces us to use the 1-Wasserstein distance instead of the other orders in the proof, even though it may be less simple to treat the diffusion terms as above.

Since the intensity of the jump processes also depends on the mean field terms, in order to treat the first term, we will use the same coupling m_s as introduced previously. One has:

$$\begin{aligned}
 I_1 &= \mathbb{E} \left[\int_0^t |K_{\mu_s^1}(X_s^1) \mathbf{1}_0(E_s^1) - K_{\mu_s^2}(X_s^2) \mathbf{1}_0(E_s^2)| ds \right] \\
 &= \mathbb{E} \left[\int_0^t \left| \int_{\Pi} K(X_s^1, x) \mathbf{1}_0(E_s^1) \mathbf{1}_1(e) \mu_s^1(dx, de) - \int_{\Pi} K(X_s^2, y) \mathbf{1}_0(E_s^2) \mathbf{1}_1(f) \mu_s^2(dy, df) \right| ds \right] \\
 &\leq \mathbb{E} \left[\int_0^t \left| \int_{\Pi \times \Pi} (K(X_s^1, x) - K(X_s^2, y)) \mathbf{1}_0(E_s^1) \mathbf{1}_1(e) m_s(dx, de, dy, df) \right| ds \right] \\
 &\quad + \mathbb{E} \left[\int_0^t \left| \int_{\Pi \times \Pi} K(X_s^2, y) (\mathbf{1}_0(E_s^1) \mathbf{1}_1(e) - \mathbf{1}_0(E_s^2) \mathbf{1}_1(f)) m_s(dx, de, dy, df) \right| ds \right] \\
 &\leq L_K \int_0^t \mathbb{E} \left[|X_s^1 - X_s^2| + \int_{\Pi \times \Pi} |x - y| m_s(dx, de, dy, df) \right] ds \\
 &\quad + \|K\|_{\infty} \int_0^t \mathbb{E} \left[|E_s^1 - E_s^2| + \int_{\Pi \times \Pi} |e - f| m_s(dx, de, dy, df) \right] ds,
 \end{aligned}$$

where to obtain the last inequality, we used the fact that $|\mathbf{1}_k(e) - \mathbf{1}_k(f)| \leq |e - f|$, for $k \in \{0, 1, 2\}$.

The same observation leads to

$$I_2 = \gamma \mathbb{E} \left[\int_0^t |\mathbf{1}_1(E_s^1) - \mathbf{1}_1(E_s^2)| ds \right] \leq \gamma \int_0^t \mathbb{E} [|E_s^1 - E_s^2|] ds.$$

Summing up the above terms, we obtain the following estimate for the changes between the epidemiological states

$$\begin{aligned}
 \mathbb{E} \left[\sup_{0 \leq s \leq t} |E_s^1 - E_s^2| \right] &\leq \mathbb{E} |E_0^1 - E_0^2| + L_K \int_0^t \mathbb{E} [|X_s^1 - X_s^2|] ds \\
 &\quad + (\gamma + \|K\|_{\infty}) \int_0^t \mathbb{E} [|E_s^1 - E_s^2|] ds \\
 &\quad + L_K \int_0^t \mathbb{E} \left[\int_{\Pi \times \Pi} |x - y| m_s(dx, de, dy, df) \right] ds \\
 &\quad + \|K\|_{\infty} \int_0^t \mathbb{E} \left[\int_{\Pi \times \Pi} |e - f| m_s(dx, de, dy, df) \right] ds.
 \end{aligned} \tag{3.2}$$

Step 3. We combine (3.1) and (3.2) to finally deduce

$$\begin{aligned}
 &\frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^1 - X_s^2| + |E_s^1 - E_s^2| \right] \\
 &\leq \mathbb{E} [|X_0^1 - X_0^2| + |E_0^1 - E_0^2|] \\
 &\quad + \left(L_V + \frac{C_1^2(L_{\sigma} + L_{\sigma_0})^2}{2} + L_K \right) \int_0^t \mathbb{E} [|X_s^1 - X_s^2|] ds \\
 &\quad + \left(L_V + \frac{C_1^2(L_{\sigma} + L_{\sigma_0})^2}{2} + \gamma + \|K\|_{\infty} \right) \int_0^t \mathbb{E} [|E_s^1 - E_s^2|] ds \\
 &\quad + (L_V + L_K) \int_0^t \mathbb{E} \left[\int_{\Pi \times \Pi} |x - y| m_s(dx, de, dy, df) \right] ds \\
 &\quad + (L_V + \|K\|_{\infty}) \int_0^t \mathbb{E} \left[\int_{\Pi \times \Pi} |e - f| m_s(dx, de, dy, df) \right] ds.
 \end{aligned}$$

Since m_s is the conditional law of the pair $((X_s^1, E_s^1), (X_s^2, E_s^2))$ given \mathcal{F}^0 , we can simplify the last two terms and deduce the following bound

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^1 - X_s^2| + |E_s^1 - E_s^2| \right] \\ & \leq 2\mathbb{E} [|X_0^1 - X_0^2| + |E_0^1 - E_0^2|] + C \int_0^t \mathbb{E} [|X_s^1 - X_s^2| + |E_s^1 - E_s^2|] ds, \end{aligned} \quad (3.3)$$

where $C = \max \left\{ (4L_V + 4L_K + C_1^2(L_\sigma + L_{\sigma_0})^2), (4L_V + 4\|K\|_\infty + 2\gamma + C_1^2(L_\sigma + L_{\sigma_0})^2) \right\}$.

Finally, we use the Gronwall lemma to get the stability estimate

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^1 - X_s^2| + |E_s^1 - E_s^2| \right] \leq 2e^{Ct} \mathbb{E} [|X_0^1 - X_0^2| + |E_0^1 - E_0^2|]. \quad (3.4)$$

Proof of i). To prove the existence of a solution to equation (2.3), we will use a fixed point theorem. We consider Θ the space consisting of Π -valued \mathcal{F} -progressively measurable *càdlàg* processes $Z = (Z_t)_{t \in [0, T]}$ on $[0, T]$ equipped with the norm

$$\|Z\|_\Theta = \mathbb{E} \left[\sup_{0 \leq t \leq T} |Z_t| \right] < \infty.$$

Notice that $Z = (X, E)$ is a mixed process, where X is continuous and E is a *càdlàg* process. With the norm defined above, $(\Theta, \|\cdot\|_\Theta)$ is a Banach space.

Now we fix $Z^1 = (X^1, E^1)$ and $Z^2 = (X^2, E^2)$ in Θ and define a mapping $\Phi : \Theta \mapsto \Theta$ which maps $(X, E) \mapsto (Y, F)$ respectively via the following stochastic integrations:

$$\begin{aligned} \forall t \in [0, T], \quad Y_t &= X_0 + \int_0^t V_{\mu_s}(X_s, E_s) ds + \int_0^t \sigma(X_s, E_s) dB_s + \int_0^t \sigma_0(X_s, E_s) dB_s^0, \\ F_t &= E_0 + \int_{[0, t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \{K_{\mu_s}(X_s) \mathbf{1}_0(E_s) + \gamma \mathbf{1}_1(E_s)\}\}} \mathcal{Q}(ds, du). \end{aligned}$$

We can show that if $(X, E) \in \Theta$, then $(Y, F) = \Phi(X, E) \in \Theta$ as well by using the Burkholder–Davis–Gundy inequality and a similar estimate as in the proof of Proposition 3.5.

If for $i = 1, 2$, $(Y^i, F^i) = \Phi(X^i, E^i)$, then by the same calculations those lead to the main estimate (3.3), we obtain

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq s \leq t} |Y_s^1 - Y_s^2| + |F_s^1 - F_s^2| \right] \\ & \leq \frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^1 - X_s^2| + |E_s^1 - E_s^2| \right] + C \int_0^t \mathbb{E} [|X_s^1 - X_s^2| + |E_s^1 - E_s^2|] ds. \end{aligned} \quad (3.5)$$

In other words, one has

$$\|\Phi(Z^1) - \Phi(Z^2)\|_\Theta \leq \left(\frac{1}{2} + C \int_0^T dt_1 \right) \|Z^1 - Z^2\|_\Theta.$$

It follows that

$$\begin{aligned}
 \|\Phi^2(Z^1) - \Phi^2(Z^2)\|_{\Theta} &\leq \frac{1}{2} \|\Phi(Z^1) - \Phi(Z^2)\|_{\Theta} + C \int_0^T \mathbb{E}[|\Phi(Z^1)_{t_1} - \Phi(Z^2)_{t_1}|] dt_1 \\
 &\leq \frac{1}{2} \left(\frac{1}{2} + C \int_0^T dt_1 \right) \|Z^1 - Z^2\|_{\Theta} + C \int_0^T \left(\frac{1}{2} + C \int_0^{t_1} dt_2 \right) \|Z^1 - Z^2\|_{\Theta} dt_1 \\
 &= \left[\left(\frac{1}{2} \right)^2 + 2 \left(\frac{1}{2} \right) C \int_0^T dt_1 + C^2 \int_0^T dt_1 \int_0^{t_1} dt_2 \right] \|Z^1 - Z^2\|_{\Theta}
 \end{aligned}$$

and in general,

$$\begin{aligned}
 \|\Phi^k(Z^1) - \Phi^k(Z^2)\|_{\Theta} &\leq \frac{1}{2} \|\Phi^{k-1}(Z^1) - \Phi^{k-1}(Z^2)\|_{\Theta} + C \int_0^T \mathbb{E}[|\Phi^{k-1}(Z^1)_{t_1} - \Phi^{k-1}(Z^2)_{t_1}|] dt_1 \\
 &\leq \dots \\
 &\leq \left[\left(\frac{1}{2} \right)^k + \binom{k}{k-1} \left(\frac{1}{2} \right)^{k-1} C \int_0^T dt_1 + \dots \right. \\
 &\quad \left. + \binom{k}{1} \left(\frac{1}{2} \right) C^{k-1} \int_0^T dt_1 \dots \int_0^{t_{k-2}} dt_{k-1} \right. \\
 &\quad \left. + C^k \int_0^T dt_1 \dots \int_0^{t_{k-2}} dt_{k-1} \int_0^{t_{k-1}} dt_k \right] \|Z^1 - Z^2\|_{\Theta}.
 \end{aligned}$$

We obtain the last bound by iterating k times the inequality (3.5). Therefore, we conclude that for any $k \in \mathbb{N}$,

$$\|\Phi^k(Z^1) - \Phi^k(Z^2)\|_{\Theta} \leq \sum_{j=0}^k \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{j!} \|Z^1 - Z^2\|_{\Theta}.$$

In the next step, we will show that

$$\sum_{j=0}^k \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{j!} \xrightarrow{k \rightarrow \infty} 0, \tag{3.6}$$

and this will imply that for k large enough, Φ^k is a contraction on Θ . Therefore, Φ has a unique fixed point.

Now, to complete the proof, we will show that the above sum (3.6) vanishes in the limit $k \rightarrow \infty$. We will use the following observations:

Let $M > 2CT$ be large enough such that for all $j > M$, one has $j! > M^j$. First we can handle the remaining part of the above sum for large k ,

$$\begin{aligned}
 \sum_{j=M+1}^k \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{j!} &< \sum_{j=M+1}^k \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{M^j} \\
 &< \sum_{j=0}^k \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{M^j} \\
 &= \left(\frac{1}{2} + \frac{CT}{M} \right)^k \xrightarrow{k \rightarrow \infty} 0.
 \end{aligned}$$

For the first $M + 1$ terms, we bound the sum as follows

$$\begin{aligned} \sum_{j=0}^M \frac{k!}{j!(k-j)!} \frac{1}{2^{k-j}} \frac{(CT)^j}{j!} &= \sum_{j=0}^M \frac{k(k-1)\dots(k-j+1)}{2^k} \frac{(2CT)^j}{(j!)^2} \\ &< \sum_{j=0}^M \frac{k^j}{2^k} \frac{(2CT)^j}{(j!)^2} \\ &< \frac{(M+1)k^M}{2^k} \max_{0 \leq j \leq M} \left\{ \frac{(2CT)^j}{(j!)^2} \right\} \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

□

Remark 3.3. *In the above proof, once we coupled two processes, we used the representation of the jumps by the Poisson Random Measures instead of the Poisson Point Processes. Even though the two representations are the same in law, the law of the couplings in these two cases are very different and the use of the Poisson Random Measures here plays an important role. With the Poisson Random Measures, most of the jumps of the coupling are done at the same time, which allows us to obtain the estimates locally uniformly in time, whereas the same does not hold for the Poisson Point Processes, which would involve the Skorohod topology instead of the uniform topology.*

Remark 3.4. *With the bound (3.5), we can always find an appropriate contraction. Indeed, to generalize the argument, let us introduce the following norm*

$$\|Z\|_{\infty, \lambda} = \sup_{0 \leq t} \left\{ e^{-\lambda t} \mathbb{E} \left[\left(\sup_{s \leq t} |Z_s| \right) \right] \right\} < \infty.$$

By (3.5), the mapping Φ satisfies

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |\Phi(Z)_s| \right] \leq \frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} |Z_s| \right] + C \int_0^t |Z_s| ds.$$

Using the norm that defined above, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq s \leq t} |\Phi(Z)_s| \right] &\leq \frac{1}{2} \mathbb{E} \left[\sup_{0 \leq s \leq t} |Z_s| \right] + C \int_0^t |Z_s| ds \\ &\leq \|Z\|_{\infty, \lambda} \left(\frac{1}{2} e^{\lambda t} + C \int_0^t e^{\lambda s} ds \right) \\ &\leq \|Z\|_{\infty, \lambda} \left(\frac{1}{2} + \frac{C}{\lambda} \right) e^{\lambda t}. \end{aligned}$$

So that

$$\|\Phi(Z)\|_{\infty, \lambda} \leq \left(\frac{1}{2} + \frac{C}{\lambda} \right) \|Z\|_{\infty, \lambda},$$

and this is a contraction if $\lambda > 2C$.

With the assumptions that the kernels and coefficients of the McKean-Vlasov equation are bounded, we have the propagation of moments up to a finite time T .

Proposition 3.5. *Assume that the initial law μ_0 has a finite moment of order q with $q \geq 1$. For all $t \leq T$, the moment of order q of μ_t is also finite.*

Proof. The proof is straightforward. First, notice that the state e is always bounded by 2, so we do not need to estimate any moment involving e . On the other hand,

$$\mathbb{E} \left[\int_{\Pi} |x|^q \mu_t(dx, de) \right] = \mathbb{E} \left[\mathbb{E} [|X_t|^q | \mathcal{F}^0] \right] = \mathbb{E} [|X_t|^q].$$

One has,

$$\begin{aligned} \mathbb{E} [|X_t|^q] &\leq C_q \mathbb{E} [|X_0|^q] + C_q \mathbb{E} \left[\left| \int_0^t V_{\mu_s}(X_s, E_s) ds \right|^q \right] \\ &\quad + C_q \mathbb{E} \left[\left| \int_0^t \sigma(X_s, E_s) dB_s \right|^q \right] + C_q \mathbb{E} \left[\left| \int_0^t \sigma_0(X_s, E_s) dB_s^0 \right|^q \right] \\ &\leq C_q \mathbb{E} [|X_0|^q] + C_q \mathbb{E} \left[\int_0^t \int_{\Pi} |V(X_s, E_s, x, e)|^q \mu_s(dx, de) ds \right] \\ &\quad + C_q \mathbb{E} \left[\left(\int_0^t |\sigma(X_s, E_s)|^2 ds \right)^{q/2} \right] + C_q \mathbb{E} \left[\left(\int_0^t |\sigma_0(X_s, E_s)|^2 ds \right)^{q/2} \right] \\ &\leq C_q \mathbb{E} [|X_0|^q] + C(q, V, \sigma, \sigma_0)T, \end{aligned}$$

where the last inequality follows from by the boundedness of V, σ, σ_0 . \square

4. THE NON LINEAR PDE

The empirical measure of the interacting system converges to the conditional law of the unique solution of the limit SDE (2.3), which is also the solution of the stochastic PDE (2.5). We will give a proof of Proposition 2.6, which shows the connection between the SPDE (2.5) and the nonlinear jumping SDE (2.3).

Proof of Proposition 2.6. Let $\phi \in C_b^2(\Pi)$ and use Ito's formula to expand

$$\begin{aligned} &\phi(X_t, E_t) - \phi(X_0, E_0) \\ &= \int_0^t \partial_x \phi(X_s, E_s) dX_s + \frac{1}{2} \int_0^t \text{Tr} [\partial_{xx}^2 \phi(X_s, E_s) d\langle X \rangle_s] \\ &\quad + \int_{[0, t] \times \mathbb{R}_+} (\phi(X_s, E_s) - \phi(X_s, E_{s-})) \mathbf{1}_{\{u \leq K_{\mu_s}(X_s) \mathbf{1}_0(E_{s-}) + \gamma \mathbf{1}_1(E_{s-})\}} Q(ds, du) \\ &= I_1 + I_2 + I_3. \end{aligned}$$

We now take the conditional expectation on the information generated up to time t by the common noise B^0 for each term above and notice that the conditional law of $\phi(X_t, E_t)$ given \mathcal{F}_t^0 is the random variable $\langle \mu_t, \phi \rangle$.

$$\begin{aligned} \mathbb{E} [I_1 | \mathcal{F}_t^0] &= \mathbb{E} \left[\int_0^t \partial_x \phi(X_s, E_s) V_{\mu_s}(X_s, E_s) ds \middle| \mathcal{F}_t^0 \right] + \mathbb{E} \left[\int_0^t \partial_x \phi(X_s, E_s) \sigma(X_s, E_s) dB_s^i \middle| \mathcal{F}_t^0 \right] \\ &\quad + \mathbb{E} \left[\int_0^t \partial_x \phi(X_s, E_s) \sigma_0(X_s, E_s) dB_s^0 \middle| \mathcal{F}_t^0 \right] \\ &= \int_0^t \int_{\Pi} \partial_x \phi(x, e) V_{\mu_s}(x, e) \mu_s(dx, de) ds + \int_0^t \int_{\Pi} \partial_x \phi(x, e) \sigma_0(x, e) \mu_s(dx, de) dB_s^0. \end{aligned}$$

In the last equality, the change of integrations w.r.t \mathcal{F}_t^0 follows a kind of Fubini's theorem for conditional expectation and stochastic integral. Indeed, we first consider the simple predictable

processes then apply the tower property since $\mathcal{F}_t^0 \subset \mathcal{F}_{\sigma(X_t)} \vee \mathcal{F}_{\sigma(E_t)}$, and the pull out property of conditional expectation, and at the end passing to limit the approximations by the dominated convergence theorem. That means, for Z_t a $(\mathcal{F}_t^0 \vee \mathcal{F}_t^{1, \dots, N})$ -predictable process, the following holds:

$$\begin{aligned} \mathbb{E} \left[\int_0^t Z_s dB_s^i | \mathcal{F}_t^0 \right] &= 0, \\ \mathbb{E} \left[\int_0^t Z_s dB_s^0 | \mathcal{F}_t^0 \right] &= \int_0^t \mathbb{E} [Z_s | \mathcal{F}_s^0] dB_s^0, \\ \mathbb{E} \left[\int_0^t Z_s ds | \mathcal{F}_t^0 \right] &= \int_0^t \mathbb{E} [Z_s | \mathcal{F}_s^0] ds = \int_0^t \mathbb{E} [Z_s | \mathcal{F}_s^0] ds. \end{aligned}$$

where in the last inequality, we used the fact that $\mathcal{F}_t^0 = \mathcal{F}_s^0 \vee \sigma \{B_r^0 - B_s^0, s \leq r \leq t\}$ and $\sigma \{B_r^0 - B_s^0, s \leq r \leq t\} \perp\!\!\!\perp Z_s$.

For the quadratic variation term we also have

$$\begin{aligned} \mathbb{E} [I_2 | \mathcal{F}_t^0] &= \mathbb{E} \left[\frac{1}{2} \int_0^t \text{Tr} [\partial_{xx}^2 \phi(X_s, E_s) (\sigma \sigma^T(X_s, E_s) + \sigma_0 \sigma_0^T(X_s, E_s))] ds \mid \mathcal{F}_t^0 \right] \\ &= \frac{1}{2} \int_0^t \int_{\Pi} \text{Tr} [\partial_{xx}^2 \phi(x, e) (\sigma \sigma^T + \sigma_0 \sigma_0^T)(x, e) \mu_s(dx, de)] ds. \end{aligned}$$

For the last term, since the counting process has a compensator

$$\int_0^s (K_{\mu_r}(X_r) \mathbb{1}_{E_{r-}=0} + \gamma \mathbb{1}_{E_{r-}=1}) dr,$$

and the integrand $(\phi(X_s, E_s) - \phi(X_s, E_{s-}))$ is left-continuous and therefore predictable, then the expectation of I_3 is equal to the conditional expectation of the integration w.r.t the compensator.

$$\begin{aligned} \mathbb{E} [I_3 | \mathcal{F}_t^0] &= \mathbb{E} \left[\int_{[0,t] \times \mathbb{R}_+} (\phi(X_s, E_s) - \phi(X_s, E_{s-})) \mathbb{1}_{\{u \leq K_{\mu_s}(X_s) \mathbb{1}_0(E_{s-}) + \gamma \mathbb{1}_1(E_{s-})\}} Q(ds, du) | \mathcal{F}_t^0 \right] \\ &= \mathbb{E} \left[\int_0^t (\phi(X_s, E_s) - \phi(X_s, E_{s-})) (K_{\mu_s}(X_s) \mathbb{1}_0(E_{s-}) + \gamma \mathbb{1}_1(E_{s-})) ds | \mathcal{F}_t^0 \right] \\ &= \int_0^t \int_{\Pi} (\phi(x, e+1) - \phi(x, e)) K_{\mu_s}(x) \mathbb{1}_{e=0} \mu_s(dx, de) ds \\ &\quad + \int_0^t \int_{\Pi} (\phi(x, e+1) - \phi(x, e)) \gamma \mathbb{1}_{e=1} \mu_s(dx, de) ds. \end{aligned}$$

Combining all the above terms, we deduce the weak form of the SPDE (2.5)

$$\begin{aligned} d \langle \phi, \mu_t \rangle &= \langle V_{\mu_t} \cdot \partial_x \phi, \mu_t \rangle dt + \frac{1}{2} \langle \text{Tr} [((\sigma \sigma^T + \sigma_0 \sigma_0^T)) \partial_{xx}^2 \phi], \mu_t \rangle dt + \langle \sigma_0 \cdot \partial_x \phi, \mu_t \rangle dB_t^0 \\ &\quad + \langle K_{\mu_t} (\mathbb{1}_{e=1} - \mathbb{1}_{e=0}) \phi, \mu_t(dx, 0) \rangle dt + \langle \gamma (\mathbb{1}_{e=2} - \mathbb{1}_{e=1}) \phi, \mu_t(dx, 1) \rangle. \end{aligned}$$

□

Remark 4.1. *The question of uniqueness of solutions to the above nonlinear SPDE is interesting, and of course related to the uniqueness of solution to the McKean-Vlasov jump-diffusion process (2.3). We can point out a strategy to prove the uniqueness of solutions to the nonlinear SPDE (2.5) in the class of measure-valued solutions relies on the uniqueness result proved for solution of the nonlinear SDE (2.3), and also the uniqueness of solutions to a corresponding linear SPDE (which is unknown so far).*

We discuss here a sketch of that strategy. Indeed, we start with a solution μ to the SPDE (2.5) and freeze the coefficients of that SPDE by setting $a(t, x) = K_{\mu_t}(x)$, $b(t, x) = V_{\mu_t}(x)$. We obtain a linear SPDE, with smooth coefficients (precisely the smoothness),

$$\begin{aligned} d\nu_t = & -\partial_x \cdot (b\nu_t) dt + \frac{1}{2} \text{Tr} [\nabla_{xx}^2 ((\sigma\sigma^T + \sigma_0\sigma_0^T)\nu_t)] dt - \partial_x \cdot (\sigma_0\nu_t) dB_t^0 \\ & + a(\mathbb{1}_{e=1} - \mathbb{1}_{e=0})\nu_t(dx, 0)dt + \gamma(\mathbb{1}_{e=2} - \mathbb{1}_{e=1})\nu_t(dx, 1)dt. \end{aligned} \tag{4.1}$$

Obviously μ solves the linear version (4.1) with coefficients a, b . If that equation has at most one solution in the class of finite measures, so that μ is the unique solution in that class to the linear SPDE (4.1).

We next consider an underlying the linear version of the McKean-Vlasov SDE with jumps (2.3), which conditional law (at fixed time) is a solution to linear SPDE (4.1) :

$$\begin{cases} dX_t &= b(X_t, E_t)dt + \sigma(X_t, E_t)dB_t + \sigma_0(X_t, E_t)dB_t^0, \\ E_t &= E_0 + P \left(\int_0^t (a(X_s)\mathbb{1}_{E_{s-}=0} + \gamma\mathbb{1}_{E_{s-}=1}) ds \right), \\ \mu_t &= \mathcal{L}((X_t, E_t)|\mathcal{F}_t^0). \end{cases} \tag{4.2}$$

Let Ω be a probability space which is rich enough to define the Brownian motions B, B^0 and the Poisson Point Process P , independent of each others, and also a r.v. Z_0 which is a \mathcal{F}_0 -measurable random variable such that $\mathcal{L}(Z_0) = \mu_0$. Using standard results in the field of SDEs, we could define a process $Z : [0, T] \times \Omega \rightarrow \Pi$ that is a \mathcal{F}_t -adapted solution to linear SDE (4.2) (with coefficients a, b). Using Itô's formula, we know that $\mathcal{L}(Z|\mathcal{F}^0)$ is also a solution to equation (4.1). Now assuming the uniqueness of solutions to the linear SPDE (4.1) in the class of finite measures, we can conclude that for any $t \geq 0$, $\mathbb{P}^0 - a.s.$, $\mu_t = \mathcal{L}(Z_t|\mathcal{F}^0)$.

But this implies that the above process Z is a solution to the limit jumping SDE (2.3). So, if we start from a second solution ν to the SPDE (2.5), with the same initial condition $\nu_0 = \mu_0$. Then, repeating the above argument, we can construct on the same probability space Ω a process Z' solution to (2.3), such that $\nu_t = \mathcal{L}(Z'_t|\mathcal{F}^0)$ for any $t \geq 0$.

Now by Theorem (2.5) which states that the uniqueness of processes solution to equation (2.3) constructed on the same probability space. This means that in fact $Z = Z'$ almost surely, and this implies that $\mathbb{P}^0 - a.s.$, $\mu = \nu$.

5. LAW OF LARGE NUMBERS AND PROPAGATION OF CHAOS

In this section, we provide a quantitative estimates for the convergence of the empirical measure of the epidemic dynamic system to the conditional law given \mathcal{F}^0 of the unique solution of the mean field limit.

We now introduce the auxiliary system as described before, which is made of independent copies of the limit SDE driven by independent noises $(B^i)_{i=1, \dots, N}$ instead of B , independent Poisson Random Measures $(\mathcal{Q}^i)_{i=1, \dots, N}$ instead of \mathcal{Q} namely

$$\begin{cases} X_t^i &= X_0^i + \int_0^t V_{\mu_s}(X_s^i, E_s^i)ds + \int_0^t \sigma(X_s^i, E_s^i)dB_s^i + \int_0^t \sigma_0(X_s^i, E_s^i)dB_s^0, \\ E_t^i &= E_0^i + \int_{[0, t] \times \mathbb{R}_+} \mathbb{1}_{\{u \leq K_{\mu_s}(X_s^i)\mathbb{1}_0(E_{s-}^i) + \gamma\mathbb{1}_1(E_{s-}^i)\}} \mathcal{Q}^i(ds, du), \\ \mu_t &= \mathcal{L}((X_t^i, E_t^i)|\mathcal{F}_t^0). \end{cases} \tag{5.1}$$

Recall that we denote by $\bar{Z}^i = (\bar{X}_t^i, \bar{E}_t^i)_{1 \leq i \leq N}$ the unique solution to that system with the initial conditions $\{(\bar{X}_0^i, \bar{E}_0^i)\}_{1 \leq i \leq N}$.

5.1. Control the distance between the original system and an auxiliary system. We can first perform a stability type estimate for the distance between the trajectories of the original system (1.3) and the system of i.i.d. agents (5.1).

Proof of Theorem 2.7. The first part of the proof is a slight modification of the proof for the stability of the solution to the SDE (2.3). Indeed, we can establish an estimate similar to (3.1) for the difference between the pairs $\{(X^{i,N}, E^{i,N}), (\bar{X}^i, \bar{E}^i)\}$, $i = 1, \dots, N$, in the path space $D([0, T], \Pi)$ endowed with the supremum in time distance. First, it is not hard to handle the diffusion terms imitating the proof of Theorem 2.5, the difference only arise at the two mean-field terms: the drift term and the jump term related to infection. Recall that conditionally on \mathcal{F}^0 , the laws of the individuals remain the same and $\mathcal{L}((\bar{X}_t^i, \bar{E}_t^i) | \mathcal{F}_t^0) = \mu_t$ for all $i \in \{1, \dots, N\}$.

We treat the drift terms first. Since the choice of the coupling m_s in the proof of Theorem 2.5 is arbitrary, at the last inequality we can obtain an appropriate bound due to the definition of the 1-Wassertein metric

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s \frac{1}{N} \sum_{j=1}^N V(X_r^{i,N}, E_r^{i,N}, X_r^{j,N}, E_r^{j,N}) dr - \int_0^s V_{\mu_r}(\bar{X}_r^i, \bar{E}_r^i) dr \right| \right] \\ & \leq \mathbb{E} \left[\int_0^t \left| V_{\mu_s^N}(X_s^{i,N}, E_s^{i,N}) - V_{\mu_s}(\bar{X}_s^i, \bar{E}_s^i) \right| ds \right] \\ & \leq C \int_0^t \mathbb{E} \left[|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| + W_1(\mu_s^N, \mu_s) \right] ds. \end{aligned}$$

For the mean-field term related to the jump, we use the same coupling and apply the same argument. To be more precise, one has

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq s \leq t} \int_{[0,s] \times \mathbb{R}_+} \left| \mathbb{1}_{\{u \leq K_r^{i,N} \mathbb{1}_0(E_r^{i,N}) + \gamma \mathbb{1}_1(E_r^{i,N})\}} - \mathbb{1}_{\{u \leq K_{\mu_r}(\bar{X}_{r-}^i) \mathbb{1}_0(\bar{E}_{r-}^i) + \gamma \mathbb{1}_1(\bar{E}_{r-}^i)\}} \right| \mathcal{Q}(dr, du) \right] \\ & \leq \mathbb{E} \left[\int_0^t \left| (K_s^{i,N} \mathbb{1}_0(E_s^{i,N}) + \gamma \mathbb{1}_1(E_s^{i,N})) - (K_{\mu_s}(\bar{X}_s^i) \mathbb{1}_0(\bar{E}_s^i) + \gamma \mathbb{1}_1(\bar{E}_s^i)) \right| ds \right] \\ & \leq \mathbb{E} \left[\int_0^t \left| K_s^{i,N} \mathbb{1}_0(E_s^{i,N}) - K_{\mu_s}(\bar{X}_s^i) \mathbb{1}_0(\bar{E}_s^i) \right| ds \right] + \gamma \mathbb{E} \left[\int_0^t \left| \mathbb{1}_1(E_s^{i,N}) - \mathbb{1}_1(\bar{E}_s^i) \right| ds \right] \\ & = I_1 + I_2. \end{aligned}$$

We first treat I_1 which contains the mean-field term $K_s^{i,N}$:

$$\begin{aligned} I_1 &= \mathbb{E} \left[\int_0^t \left| \int_{\Pi} K(X_s^{i,N}, x) \mathbb{1}_0(E_s^{i,N}) \mathbb{1}_1(e) \mu_s^N(dx, de) - \int_{\Pi} K(\bar{X}_s^i, y) \mathbb{1}_0(\bar{E}_s^i) \mathbb{1}_1(f) \mu_s(dy, df) \right| ds \right] \\ & \leq C \int_0^t \mathbb{E} \left[|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| + W_1(\mu_s^N, \mu_s) \right] ds. \end{aligned}$$

The remaining terms are treated similarly as in the proof of Theorem 2.5. Combining them we can deduce the following bound:

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| \right] \\ & \leq 2\mathbb{E} \left[|X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i| \right] + C \int_0^t \mathbb{E} \left[|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| + W_1(\mu_s^N, \mu_s) \right] ds. \end{aligned}$$

On the other hand, one has

$$\begin{aligned}
 \mathbb{E} [W_1 (\mu_s^N, \mu_s)] &= \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{Z_s^{i,N}}, \mu_s \right) \right] \\
 &\leq \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{Z_s^{i,N}}, \frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i} \right) \right] + \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i}, \mu_s \right) \right] \\
 &\leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i|] + \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i}, \mu_s \right) \right].
 \end{aligned} \tag{5.2}$$

From this and the fact that $(Z^{i,N}, \bar{Z}^i), i = 1, \dots, N$ are identically distributed, we have the following

$$\begin{aligned}
 &\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| \right] \\
 &\leq 2\mathbb{E} [|X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i|] \\
 &\quad + C \int_0^t \mathbb{E} \left[|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| + W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i}, \mu_s \right) \right] ds.
 \end{aligned}$$

Now applying the Gronwall lemma, we deduce that

$$\begin{aligned}
 &\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| \right] \\
 &\leq e^{Ct} \left(2\mathbb{E} [|X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i|] + \int_0^t \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i}, \mu_s \right) \right] ds \right).
 \end{aligned}$$

Since conditionally upon \mathcal{F}^0 , $\{\bar{Z}_t^i = (\bar{X}_t^i, \bar{E}_t^i)\}_{i=1, \dots, N}$ are N i.i.d.r.v., by the Glivenko-Cantelli theorem, their empirical measure converges *a.s* to μ_t . So the second term in the r.h.s. in the above bound goes to 0. Moreover, if $\mathbb{E} [|Z_t^i|^q] < \infty$ for some $q > 2$, it is possible to give a rate of convergence measured in the Wasserstein distance. Indeed, we are able to adapt to the conditional case the crucial result obtained by Fournier-Guillin [12]. See Proposition 6.3 in the Appendix for a precise statement.

Hence by Propositions 6.3 and 3.5, we can deduce the desired stability estimate

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i| \right] \leq C(t) \left(\mathbb{E} [|X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i|] + \alpha_d(N) \right), \tag{5.3}$$

with $\alpha_d(N)$ as defined in the statement of Theorem 2.7. □

As a consequence, at any time t , we can give a quantitative estimate of convergence of the empirical measure of the original system towards its expected limit μ_t , measured in the Wasserstein distance.

5.2. Proof of Corollary 2.8.

Proof. Combining (5.2) and (5.3), we obtain

$$\begin{aligned} \mathbb{E} [W_1(\mu_s^N, \mu_s)] &\leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|X_s^{i,N} - \bar{X}_s^i| + |E_s^{i,N} - \bar{E}_s^i|] + \mathbb{E} \left[W_1 \left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{Z}_s^i}, \mu_s \right) \right] \\ &\leq C(t) \left(\mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N |X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i| \right] + \alpha_d(N) \right). \end{aligned}$$

Thanks to the exchangeability and a coupling result, see for instance [13, Proposition A.1] or [15, Proposition 2.14], we have the right to choose the initial conditions of the coupling such that, almost surely,

$$W_1(\mu_0^N, \bar{\mu}_0^N) = \frac{1}{N} \sum_{i=1}^N |X_0^{i,N} - \bar{X}_0^i| + |E_0^{i,N} - \bar{E}_0^i|,$$

which allows us to deduce

$$\begin{aligned} \mathbb{E} [W_1(\mu_t^N, \mu_t)] &\leq C(t) (\mathbb{E} [W_1(\mu_0^N, \bar{\mu}_0^N)] + \alpha_d(N)) \\ &\leq C(t) (\mathbb{E} [W_1(\mu_0^N, \mu_0)] + \mathbb{E} [W_1(\bar{\mu}_0^N, \mu_0)] + \alpha_d(N)) \\ &\leq C(t) (\mathbb{E} [W_1(\mu_0^N, \mu_0)] + \alpha_d(N)), \end{aligned}$$

where we allow the constant to change from line to line, and use the quantitative bound for $W_1(\bar{\mu}_0^N, \mu_0)$ for the last inequality. \square

6. APPENDIX

6.1. Definitions of Poisson Point Process and Poisson Random Measure.

Definition 6.1. An \mathbb{N} -valued process $(P_t)_{t \geq 0}$ is said to be a standard Poisson Point Process if $P_0 = 0$, its increments over disjoint intervals are independent, and for $0 \leq s \leq t$, $P_t - P_s$ follows the $\text{Poi}(t - s)$ distribution. This implies in particular that all the jumps of P_t have size 1.

Definition 6.2. A measure Q on \mathbb{R}_+^2 is said to be a standard Poisson Random Measure if it is a sum of Dirac measures located at countably many points of \mathbb{R}_+^2 , in such a way that the numbers of those points in disjoint Borel sets are mutually independent, and the number of points in a Borel subset $A \subset \mathbb{R}_+^2$ follows the $\text{Poi}(\mu(A))$ distribution, where μ denotes the Lebesgue measure on \mathbb{R}_+^2 .

The mean measure of a standard Poisson Random Measure is the Lebesgue measure. Non standard Poisson Random Measures can have more general mean measures. The same is true for non standard Poisson Point Processes.

6.2. Proof of the equivalences in Definition 2.3.

Proof. (i) \Rightarrow (ii) is obvious.

(ii) \Rightarrow (iii). Taking any $\phi \in C_b(\Pi)$,

$$\begin{aligned} \mathbb{E} \left[\langle \mu^N - \mu, \phi \rangle^2 \mid \mathcal{F}^0 \right] &= \frac{1}{N^2} \sum_{i,j=1}^N \mathbb{E} [\phi(Z^{i,N}) \phi(Z^{j,N}) \mid \mathcal{F}^0] \\ &\quad - \frac{2}{N} \sum_{i=1}^N \mathbb{E} [\phi(Z^{i,N}) \langle \mu, \phi \rangle \mid \mathcal{F}^0] + \mathbb{E} [\langle \mu, \phi \rangle^2 \mid \mathcal{F}^0]. \end{aligned}$$

Using the conditional exchangeability we rewrite the r.h.s. as

$$\frac{1}{N} \mathbb{E} [\phi(Z^{1,N})^2 | \mathcal{F}^0] + \frac{N-1}{N} \mathbb{E} [\phi(Z^{1,N})\phi(Z^{2,N}) | \mathcal{F}^0] - 2 \langle \mu, \phi \rangle \mathbb{E} [\phi(Z^{1,N}) | \mathcal{F}^0] + \langle \mu, \phi \rangle^2,$$

which tends to 0 by (ii). Therefore,

$$\mathbb{E} [\langle \mu^N, \phi \rangle | \mathcal{F}^0] \rightarrow \langle \mu, \phi \rangle.$$

Since μ is \mathcal{F}^0 -measurable and the above statement is true for any bounded continuous ϕ , we conclude that μ^N converges weakly to μ \mathbb{P}^0 -a.s.

(iii) \Rightarrow (i). We consider only the case $k = 2$ of that implication, but the general case $k \geq 2$ could be handled in a similar way.

Let $\phi_1, \phi_2 \in C_b(\Pi)$. We will prove the convergence in law of $(Z^{1,N}, Z^{2,N})$ using only functions of the form $\phi(z_1, z_2) = \phi_1(z_1)\phi_2(z_2)$, whose linear combinations are dense in $C_p(\Pi^2)$. By the triangular inequality, one has

$$\begin{aligned} & \left| \mathbb{E} [\phi_1(Z^{1,N})\phi_2(Z^{2,N}) | \mathcal{F}^0] - \langle \mu, \phi_1 \rangle \langle \mu, \phi_2 \rangle \right| \\ & \leq \left| \mathbb{E} [\phi_1(Z^{1,N})\phi_2(Z^{2,N}) | \mathcal{F}^0] - \mathbb{E} [\langle \mu^N, \phi_1 \rangle \langle \mu^N, \phi_2 \rangle | \mathcal{F}^0] \right| \end{aligned} \quad (6.1)$$

$$+ \left| \mathbb{E} [\langle \mu^N, \phi_1 \rangle \langle \mu^N, \phi_2 \rangle | \mathcal{F}^0] - \mathbb{E} [\langle \mu, \phi_1 \rangle \langle \mu, \phi_2 \rangle | \mathcal{F}^0] \right|. \quad (6.2)$$

Using the conditional exchangeability, we have an upper bound for (6.1),

$$\begin{aligned} & \left| \mathbb{E} [\phi_1(Z^{1,N})\phi_2(Z^{2,N}) | \mathcal{F}^0] - \mathbb{E} [\langle \mu^N, \phi_1 \rangle \langle \mu^N, \phi_2 \rangle | \mathcal{F}^0] \right| \\ & = \left| \frac{1}{N(N-1)} \sum_{i,j=1, i \neq j}^N \mathbb{E} [\phi_1(Z^{i,N})\phi_2(Z^{j,N}) | \mathcal{F}^0] - \frac{1}{N^2} \sum_{i,j=1}^N \mathbb{E} [\phi_1(Z^{i,N})\phi_2(Z^{j,N}) | \mathcal{F}^0] \right| \\ & \leq \left| \left(\frac{1}{N(N-1)} - \frac{1}{N^2} \right) \sum_{i,j=1, i \neq j}^N \mathbb{E} [\phi_1(Z^{i,N})\phi_2(Z^{j,N}) | \mathcal{F}^0] \right| \\ & \quad + \left| \frac{1}{N^2} \sum_{i=1}^N \mathbb{E} [\phi_1(Z^{i,N})\phi_2(Z^{i,N}) | \mathcal{F}^0] \right| \\ & \leq \frac{1}{N} \|\phi_1\|_\infty \|\phi_2\|_\infty + \frac{1}{N} \|\phi_1\|_\infty \|\phi_2\|_\infty, \end{aligned}$$

where the last quantity tends to 0 as $N \rightarrow \infty$.

Considering (6.2), we can prove that it converges to 0, using the point (iii) and the fact that the function $\tilde{\mu} \mapsto \langle \tilde{\mu}, \phi_1 \rangle \langle \tilde{\mu}, \phi_2 \rangle$ is continuous on $\mathcal{P}(\Pi)$. □

6.3. Proof of Lemma 2.4.

Proof. The system of equation (1.3) is symmetric under any permutation of the individuals and their individual noises. Hence, if $Z_t^N := (Z_t^1, \dots, Z_t^N)$ is a strong solution of the system with initial condition (Z_0^1, \dots, Z_0^N) then $Z_t^{\sigma, N} := (Z_t^{\sigma(1), N}, \dots, Z_t^{\sigma(N), N})$ is also a strong solution of the equation with initial condition $(Z_0^{\sigma(1)}, \dots, Z_0^{\sigma(N)})$ with permuted individual noise. Moreover, by the assumption that the initial laws of Z_0^N and $Z_0^{\sigma, N}$ are equal, the conclusion follows using

the weak uniqueness of solutions to the SDE system (1.3) with jumps. In fact, the standard approach of [16, Theorem 9.1] applies to our particular case, since all the coefficients appearing in the system (1.3) are regular enough. \square

6.4. Rate of convergence in the empirical law of large number: the conditionally independence case.

Proposition 6.3. *For any $d \in \mathbb{N}$ and $q \in (2, \infty)$, there exists a constant $C_{d,q}$ such that the following holds. For any \mathbb{P}_0 -measurable random probability μ on Π satisfying*

$$\mathbb{E} \left[\int_{\mathbb{R}^d} |z|^q \mu(dz) \right] = \mathbb{E}[|Z_1|^q] < \infty,$$

if conditionally upon \mathcal{F}_0 , the Π -valued random variables $(Z_i)_{i \geq 1}$ are i.i.d. with conditional law μ , and denoting the random empirical measure by $\mu_Z^N = \frac{1}{N} \sum_{i=1}^N \delta_{Z_i}$, we have

$$\mathbb{E} [W_1(\mu_Z^N, \mu)] \leq C_{d,q} \mathbb{E} [|Z_1|^q]^{1/q} \begin{cases} N^{-1/2}, & d = 1, \\ N^{-1/2} \log N, & d = 2, \\ N^{-1/d}, & d \geq 3. \end{cases}$$

Proof. For a given realization of the common noise, we rely first on a result by Fournier Guillin [12, Theorem 1] (which is the best result so far after a long sequence of previous partial or less accurate result) in the case of i.i.d. random variables with a given (deterministic) common law μ . Applying it (with $p = 1$) in our special case⁴, it gives the following estimate conditionally on \mathcal{F}^0 ,

$$\mathbb{P}^0 - \text{a.s.}, \quad \mathbb{E} [W_1(\mu_Z^N, \mu) | \mathcal{F}^0] \leq C_{d,q} \mathbb{E} [|Z_1|^q | \mathcal{F}^0]^{1/q} \begin{cases} N^{-1/2} + N^{-(q-1)/q}, & d = 1, \\ N^{-1/2} \log N + N^{-(q-1)/q}, & d = 2, \\ N^{-1/d} + N^{-(q-1)/q}, & d \geq 3. \end{cases}$$

First, we could remove all the second power of N involving q in the three cases since it is always smaller than the first term when $q > 2$.

Taking then the expectation w.r.t. to \mathbb{P}_0 in the above inequality, and using (since $q > 1$) the Jensen inequality on the q -th moment

$$\mathbb{E} \left[\mathbb{E} [|Z_1|^q | \mathcal{F}^0]^{1/q} \right] \leq \mathbb{E} [|Z_1|^q]^{1/q},$$

we get the desired result. \square

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability

The document contains no data.

REFERENCES

- [1] P. Billingsley (1968). *Convergence of Probability Measures*, John Wiley & Sons, New York.
- [2] O. N. Bjørnstad, B. T. Grenfell, C. Viboud, A. A. King (2019). Comparison of alternative models of human movement and the spread of disease, bioRxiv 2019.12.19.882175; doi: <https://doi.org/10.1101/2019.12.19.882175>
- [3] S. Bowong, A. Emakoua, E. Pardoux (2022). A Spatial Stochastic Epidemic Model: Law of Large Numbers and Central Limit Theorem, *Stochastic and Partial Diff. Equations*, to appear. arXiv:2007.06636.

⁴remark that the particular assumption on q in [12, Theorem 1] are fulfilled by our particular choice $q > 2$

- [4] T. Britton, E. Pardoux (2019). Stochastic epidemic in a homogeneous community, Part I of *Stochastic epidemic models with inference*, Lecture Notes in Mathematics **2225**, pp. 1-120 Springer.
- [5] D. Brockmann, L. Hufnagel, Geisel (2006). The scaling laws of human travel. *Nature* **439**, 462–465 <https://doi.org/10.1038/nature04292>
- [6] J.A. Carrillo, Y. P. Choi, M. Hauray (2014). The derivation of swarming models: Mean-field limit and Wasserstein distances. In: Muntean A., Toschi F. (eds) *Collective Dynamics from Bacteria to Crowds*. CISM International Centre for Mechanical Sciences (553), Springer. <https://arxiv.org/abs/1304.5776>
- [7] R. Carmona, and F. Delarue (2018). *Probabilistic theory of mean field games with applications. II*, Probability Theory and Stochastic Modelling **84**, Springer.
- [8] R. Carmona, F. Delarue and D. Lacker (2016). Mean field games with common noise, *The Annals of Probability* **44**, 3740–3803.
- [9] M. Coghi, F. Flandoli (2016). Propagation of chaos for interacting particles subject to environmental noise, *The Annals of Applied Probability*, **26**, No. 3, 1407–1442.
- [10] S. Erlander and N. F. Stewart, (1990). *The Gravity Model in Transportation Analysis: Theory and Extensions*, CRC Press.
- [11] X. Erny, E. Löcherbach, D. Loukianova (2020). Conditional propagation of chaos for mean field systems of interacting neurons, arXiv:1909.02925v3.
- [12] N. Fournier and A. Guillin (2015). On the rate of convergence in Wasserstein distance of the empirical measure, *Probability Theory Related Fields* **162**, 707–738.
- [13] N. Fournier, M. Hauray (2016). Propagation of chaos for the Landau equation with moderately soft potentials. *The Annals of Probability* **44**, 3581–3660.
- [14] M. Hauray and S. Salem (2019). Propagation of chaos for the Vlasov-Poisson-Fokker-Planck system in 1D, *Kinetic and Related Models* **12**, Issue 2, pp 269-302.
- [15] M. Hauray and S. Mischler (2014). On Kac’s chaos and related problems, *J. Funct. Anal.* **266**, 6055–6157.
- [16] N. Ikeda and S. Watanabe (1989). *Stochastic Differential Equations and Diffusion Processes*, Second Edition. North-Holland Mathematical Library 24, Amsterdam.
- [17] T. Kurtz, J. Xiong (1999). Particle representations for a class of nonlinear SPDEs, *Stochastic Processes and their Applications* **83**, 103–126.
- [18] C. Leonard (1990). Some epidemic systems are long range interacting particle systems, in *Stochastic Processes in Epidemic Theory*, J.P. Gabriel et al., eds., Springer Verlag.
- [19] T. S. Lim, Y. Lu, J. Nolen (2019). Quantitative Propagation of Chaos in the bimolecular chemical reaction-diffusion model, arXiv:1906.01051v2.
- [20] H. P. McKean (1967). Propagation of chaos for a class of non-linear parabolic equations, In *Stochastic Differential Equations*, Lecture Series in Differential Equations (7), Catholic Univ., 41–57.
- [21] L. Roques, O. Bonnefon, V. Baudrot, S. Soubeyrand, H. Berestycki (2020). A parsimonious model for spatial transmission and heterogeneity in the Covid-19 propagation, *Royal Soc. Open Sci.* **7**: 201382.
- [22] M. Schlöpfer, L. Dong, K. O’Keeffe, et al. (2021). The universal visitation law of human mobility, *Nature* **593**, 522–527, <https://doi.org/10.1038/s41586-021-03480-9>
- [23] Simini, F., González, M., Maritan, A. et al. (2012). A universal model for mobility and migration patterns, *Nature* **484**, 96–100, <https://doi.org/10.1038/nature10856>
- [24] C. Song, T. Koren, P. Wang, et al. (2010). Modeling the scaling properties of human mobility, *Nature Phys.* **6**, 818–823, <https://doi.org/10.1038/nphys1760>
- [25] A-S. Sznitman (1991). Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math*, 165–251, Springer, Berlin.
- [26] C. Villani (2003). *Topics in optimal transportation*, Graduate Studies in Mathematics **58**, American Mathematical Society, Providence, RI.