



HAL
open science

A note on Fokker-Planck equations and graphons

Fabio Coppini

► **To cite this version:**

| Fabio Coppini. A note on Fokker-Planck equations and graphons. 2021. hal-03138054v1

HAL Id: hal-03138054

<https://hal.science/hal-03138054v1>

Preprint submitted on 10 Feb 2021 (v1), last revised 19 Mar 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A NOTE ON FOKKER-PLANCK EQUATIONS AND GRAPHONS

FABIO COPPINI

ABSTRACT. Fokker-Planck equations represent a suitable description of the finite-time behavior for a large class of particle systems as the size of the population tends to infinity. Recently, the theory of graph limits have been introduced in the mean-field framework to account for heterogeneous interactions among particles. In many instances, such network heterogeneity is preserved in the limit which turns from being a single Fokker-Planck equation (also known as McKean-Vlasov) to an infinite system of non-linear partial differential equations (PDE) coupled by means of a graphon. While appealing from an applied viewpoint, few rigorous results exist on the graphon particle system. This note addresses such limit systems focusing on the relation between initial conditions and interaction network: if the system initial datum and the graphon degrees satisfy a suitable condition, a significantly simpler representation of the solution is available. This in turn implies that very different graphons can lead to exactly the same particle behavior, shedding some light on the network influence on the dynamics.

Examples of such representation are provided. In particular, step kernels represent a class of graphons to which our result applies: this in turn opens the way to approximate the graphon particle system with a finite system of Fokker-Planck equations. As a byproduct, we show that when the initial condition is uniform, every graphon with constant degree leads to a behavior indistinguishable from the well-known mean-field limit.

2020 MSC: 35Q84, 05C82, 82B20, 60H20, 60K35, 35Q70.

Keywords: Fokker-Planck equation, graphons, mean-field systems, Interacting particles, McKean-Vlasov, graphon particle system, step kernels.

1. INTRODUCTION, MODEL AND LITERATURE

1.1. Introduction, aim of this note and organization. In the last years, the study of interacting particle systems with a non-trivial dense network structure has been repeatedly addressed in the mathematical community: see, e.g., [4, 10, 21, 9] for interacting diffusions, [5, 6, 20] for applications in mean-field games and [11] in the context of dynamical systems. Depending on the setting, many results on interacting particle systems are nowadays available [1, 3, 8, 7, 17] whenever the underlying graph sequence is converging, in some sense depending case by case, to a suitable object. More precisely, if the graph limit is a graphon, then, as the size of the system tends to infinity, the finite-time population behavior is suitably described by an infinite system of coupled non-linear Fokker-Planck equations, the coupling between equations being made by the graphon limit itself (see equation (1.1) for a simple example).

This note addresses the graphon particle system obtained in the limit and tries to make clear how potentially different graphons can lead to the same particle behavior. Such a perspective is important not only to better understand the limit system, but also in case one is interested in reconstructing the network structure by looking at the particle dynamics: we prove that, depending on the initial conditions, the class of suitable graphons

can be potentially very large. As it will be clear later on, the relation between the graphon and the initial datum plays the most important role.

Graphons have been used as a model for many real-world networks, yet, to the author's knowledge, known results on graphon particle systems are limited to existence and uniqueness of solutions; we refer to Subsection 1.3 for the current literature. A mathematical study of the limit object may lead to a better understanding of the complex phenomena at the heart of these models.

A labeled graphon (we use the same notation of [15]) is a symmetric measurable function W defined on the unit-square

$$\begin{aligned} W : [0, 1]^2 &\rightarrow [0, 1] \\ (x, y) &\mapsto W(x, y). \end{aligned}$$

If the limit population is represented as a continuum of particles labeled by the unit interval $[0, 1]$, then $W(x, y)$ stands for the connection strength between the particle labeled with x and the one labeled with y . The function W thus describes the connection network underlying a (possibly infinite) particle population.

Fix a finite time horizon $T > 0$. For one-dimensional particles which are interacting through a (regular enough) function Γ , the graphon particle system is an infinite system of partial differential equations (PDE) coupled by means of W , i.e.,

$$\begin{aligned} \partial_t \mu_t(\theta, x) &= \frac{1}{2} \partial_\theta^2 \mu_t(\theta, x) \\ &\quad - \partial_\theta \left[\int_0^1 W(x, y) \mu_t(\theta, x) \int_{\mathbb{R}} \Gamma(\theta, \theta') \mu_t(d\theta', y) dy \right], \quad x \in [0, 1], \end{aligned} \tag{1.1}$$

for $t \in [0, T]$ and where $\mu = \{\mu_t(\cdot, x), t \in [0, T]\}_{x \in [0, 1]}$ is a collection of probability measures. The initial datum is given by a probability measure $\mu_0 \in \mathcal{P}(\mathbb{R} \times [0, 1])$. From the disintegration theorem ([13]), we have that $\mu_0(\cdot, x) \in \mathcal{P}(\mathbb{R})$ for almost every $x \in [0, 1]$. Namely, the probability measure

$$\mu(x) := \mu(\cdot, x) = \{\mu_t(\cdot, x), t \in [0, T]\} \in \mathcal{P}(C([0, T], \mathbb{R}))$$

represents the law of the trajectory associated to the x -labeled particle, this last one being connected to the others by means of $W(x, \cdot) : [0, 1] \rightarrow [0, 1]$, see the following (1.2). Each trajectory is a continuous function from $[0, T]$ to \mathbb{R} , i.e., an element of the space $C([0, T], \mathbb{R})$.

To the system of non-linear Fokker-Planck equations (1.1), it is associated a family of continuous processes $\{\theta^x\}_{x \in [0, 1]} \subset C([0, T], \mathbb{R})$, which solve

$$\begin{cases} \theta_t^x = \theta_0^x + \int_0^t \int_0^1 W(x, y) \int_{\mathbb{R}} \Gamma(\theta_s^x, \theta) \mu_s(d\theta, y) dy ds + B_t^x, & t \in [0, T], \\ \mu_t(\cdot, y) = \mathcal{L}(\theta_t^y), & \text{for } t \in [0, T] \text{ and } y \in [0, 1], \end{cases} \tag{1.2}$$

where the law of the initial condition θ_0^x , denoted by $\mathcal{L}(\theta_0^x)$, is thus given by $\mu_0(x)$. The family $\{B^x\}_{x \in [0, 1]}$ is composed of independent and identically distributed (IID) Brownian motions, independent of the initial conditions as well.

The link between (1.1) and (1.2) is given by the fact that $\mu_t(x) = \mathcal{L}(\theta_t^x)$ for every $t \in [0, T]$ and $x \in [0, 1]$. We refer to Proposition 1.3 for a precise statement which includes more general systems, including the degenerate case ($\sigma \equiv 0$).

System (1.1) or, equivalently, the family of non-linear processes (1.2), have been proposed as a limit description in the literature [17, 21, 5, 1, 3], yet very little is known on

their mathematical properties. We thus aim at mathematically addressing (1.1), firstly by showing the strong link with the graphon theory (including unlabeled graphons, see Proposition 1.5), secondly by proving that a simpler representation for both (1.1) and (1.2) exists, whenever the initial conditions and the underlying graphon satisfy a suitable condition, see Theorem 2.1 and, in particular, Corollary 2.3 and Proposition 2.6.

The note is organized as follows: the next subsection presents the general model associated to (1.2), together with known results from the existing literature. Related works are discussed at the end of this first section.

Section 2 provides the main result, Theorem 2.1, as well as two corollaries and a relevant application to step kernels, see Proposition 2.6. Notably, Corollary 2.4 addresses the classical mean-field scenario showing that, if every particle has the same initial law, then for every graphon with constant degree, i.e., such that $d(\cdot) = \int_0^1 W(\cdot, y)dy$ is constant, the dynamics is mean-field.

Section 3 contains the mathematical set-up and the proof of the main result.

1.2. The model and some known results. We consider a class of models slightly more general than (1.2): fix $\mu_0 \in \mathcal{P}(\mathbb{R} \times [0, 1])$ and let $\{\theta^x\}_{x \in [0, 1]}$ be the family solving the ∞ -dimensional coupled system:

$$\begin{cases} \theta_t^x = \theta_0^x + \int_0^t F(\theta_s^x)ds + \int_0^t \int_0^1 W(x, y) \int_{\mathbb{R}} \Gamma(\theta_s^x, \theta) \mu_s(d\theta, y) dy ds + \int_0^t \sigma(\theta_s^x) dB_s^x \\ \mu_t(\cdot, y) = \mathcal{L}(\theta_t^y), \quad \text{for } t \in [0, T] \text{ and } y \in [0, 1], \end{cases} \quad (1.3)$$

where F , Γ and σ are 1-Lipschitz functions bounded by 1, and $\{B^x\}_{x \in [0, 1]}$ is a family of IID Brownian motions on \mathbb{R} . For every $x \in [0, 1]$, the initial condition θ_0^x is a random variable with law given by $\mathcal{L}(\theta_0^x) = \mu_0(x)$, independent of the other initial conditions and of the Brownian motions. The probability measure induced by the initial conditions and the family of Brownian motions is denoted by \mathbf{P} and the corresponding expectation by \mathbf{E} .

We work under the following assumptions.

Hypothesis 1.1. *We assume that*

- (1) (*measurability*) *The map $[0, 1] \ni x \mapsto \mu_0(x) \in \mathcal{P}(\mathbb{R})$ is measurable;*
- (2) (*moment condition*) *$\mathbf{E}[|\theta_0^x(0)|^2] < \infty$ for every $x \in [0, 1]$.*

Remark 1.2. *Measurability with respect to x is somehow necessary since we are working with measurable functions W on the unit interval $[0, 1]$; we will explicitly use it in the following relation (2.1).*

The finite second moment condition is related to the use of the 2-Wasserstein distance between probability measures (see the Section 3 and, in particular, (3.1)). It is obviously possible to work with initial conditions with finite p -moment for every $p \geq 1$, and the choice $p = 2$ is purely arbitrary.

We stick to the one-dimensional setting, i.e., θ^x taking values in \mathbb{R} ; however, all the proofs presented below are easily extendable to any finite dimension.

Existence and uniqueness for system (1.3) are known.

Proposition 1.3. *Under the measurability assumption and the moment condition in Hypothesis 1.1, there exists a unique pathwise solution to system (1.3).*

Moreover, if we denote by $\mu(x)$ the law of θ^x for each $x \in [0, 1]$, then the map $[0, 1] \ni x \mapsto \mu(x) \in \mathcal{P}(C([0, T], \mathbb{R}))$ is measurable and $\mu(x)$ weakly solves

$$\begin{aligned} \partial_t \mu_t(\theta, x) &= \frac{1}{2} \partial_\theta^2 (\sigma^2(\theta) \mu_t(\theta, x)) - \partial_\theta (\mu_t(\theta, x) F(\theta)) \\ &\quad - \partial_\theta \left[\int_0^1 W(x, y) \mu_t(\theta, x) \int_{\mathbb{R}} \Gamma(\theta, \theta') \mu_t(d\theta', y) dy \right], \end{aligned} \quad (1.4)$$

for every $x \in [0, 1]$.

We refer to [1, Proposition 2.1] and [17, Proposition 2.4] for two standard similar proofs. We point out that W needs only to be bounded and not necessarily with values in $[0, 1]$, i.e, existence and uniqueness hold for every kernel W , as in the notation below.

Let $\mathcal{W} := \{W : [0, 1]^2 \rightarrow \mathbb{R} \text{ bounded, symmetric and measurable}\}$ be the space of kernels¹. The cut-norm of $W \in \mathcal{W}$ is defined as

$$\|W\|_{\square} := \max_{S, T \subset [0, 1]} \left| \int_{S \times T} W(x, y) dx dy \right| \quad (1.5)$$

where the maximum is taken over all measurable subsets S and T of I . Let $\mathcal{W}_0 := \{W \in \mathcal{W} : 0 \leq W \leq 1\}$ be the space of labeled graphons: for $W, V \in \mathcal{W}_0$ their cut-distance is defined by

$$\delta_{\square}(W, V) := \min_{\varphi \in S_{[0, 1]}} \|W - V^{\varphi}\|_{\square}, \quad (1.6)$$

where the minimum ranges over $S_{[0, 1]}$ the space of invertible measure preserving maps from $[0, 1]$ into itself and where $V^{\varphi}(x, y) := V(\varphi(x), \varphi(y))$ for $x, y \in [0, 1]$.

The cut-distance δ_{\square} is a pseudometric on \mathcal{W}_0 since it can be zero between two different labeled graphons. If we identify all labeled graphons with cut-distance zero, we obtain the space of graphons $\widetilde{\mathcal{W}}_0 := \mathcal{W}_0 / \delta_{\square}$. A well-known result of graph limits theory says that $(\widetilde{\mathcal{W}}_0, \delta_{\square})$ is a compact metric space [15, Theorem 9.23].

Proposition 1.3 can thus be restated by saying that for every kernel $W \in \mathcal{W}$, there exists a unique solution μ^W to (1.4). It is thus natural to ask whether the application $W \mapsto \mu^W$ is continuous (e.g., in the topology of the weak-convergence) with respect to $\|\cdot\|_{\square}$. This point is discussed in the next remark.

Remark 1.4. *Under suitable assumptions on the coefficients, see [3, Proposition 3.3] but also [1, Theorem 2.1] and [14, Lemma 2.7], it is possible to show that for two solutions μ^W and μ^V associated to W and V in \mathcal{W} respectively, it holds that*

$$D_T^2(\mu^W, \mu^V) \leq C \|W - V\|_{\square}, \quad \text{for some } C > 0, \quad (1.7)$$

where D_T is some distance on probability measures metricizing the weak-convergence. Equation (1.7) proves that the application $W \mapsto \mu^W$ is (Hölder-)continuous and, thus, that similar graphons (in cut-norm) leads to similar particle behaviors. However, it does not say anything whether two different graphons (possibly in δ_{\square} -distance) lead to similar behaviors. As Theorem 2.1 and its corollaries show, there are relevant examples where this happens and where it is possible to prove that the particle dynamics coincide.

For every $W \in \mathcal{W}_0$, Proposition 1.3 shows that the infinite system (1.4) admits a unique solution, yet it does not address the unlabeled class of W in $\widetilde{\mathcal{W}}_0$. We explicit such relation in the next proposition.

¹we always consider two kernels to be equal if and only if they differ on a subset of Lebesgue measure zero. We follow closely the notation in [15].

Proposition 1.5. *Let U be a uniform random variable on $[0, 1]$ and θ^U the non-linear process in $\{\theta^x\}_{x \in [0,1]}$ with random label U . Then, the law of θ^U is given by $\tilde{\mu} = \int_0^1 \mu(x) dx$ and it is independent of the class of W in $\widetilde{\mathcal{W}}_0$.*

Proof. This result was firstly stated in [3, Proposition 2.1] in the case of interacting oscillators. In case the particles are living in \mathbb{R} the proof does not change. Indeed, let φ be a measure preserving map from $[0, 1]$ to itself, we observe that $\mu(\varphi(x))$ solves:

$$\begin{aligned} \partial_t \mu_t(\theta, \varphi(x)) &= \frac{1}{2} \partial_\theta^2 (\sigma^2(\theta) \mu_t(\theta, \varphi(x))) - \partial_\theta (\mu_t(\theta, \varphi(x)) F(\theta)) \\ &\quad - \partial_\theta \left[\int_0^1 W(\varphi(x), y) \mu_t(\theta, \varphi(x)) \int_{\mathbb{R}} \Gamma(\theta, \theta') \mu_t(d\theta', y) dy \right], \end{aligned} \quad (1.8)$$

where the last term is equal to

$$\partial_\theta \left[\int_0^1 V(x, y) \mu_t(\theta, \varphi(x)) \int_{\mathbb{R}} \Gamma(\theta, \theta') \mu_t(d\theta', \varphi(y)) dy \right]$$

with $V(x, y) = W(\varphi(x), \varphi(y))$. Thus $\{\mu(\varphi(x))\}_{x \in [0,1]}$ solves the same system of $\{\nu(x)\}_{x \in [0,1]}$ solution to (1.4) with $V \in \mathcal{W}_0$ and initial condition $\{\mu_0(\varphi(x))\}_{x \in [0,1]}$. Clearly $\tilde{\mu} = \int_0^1 \mu(x) dx = \int_0^1 \mu(\varphi(x)) dx = \int_0^1 \nu(x) dx = \tilde{\nu}$. \square

Proposition 1.5 makes clear that, as for a graphon the relevant information is independent of the labeling, the same holds true for the behavior of an interacting particle system: it does not change under relabeling of the particles. Surprisingly, this result has never been stated in the current literature (with the exception of [3] for interacting oscillators).

1.3. Related works. The graphon framework [16, 15] has been introduced in the theory of particle systems [7, 18, 19, 21] as an useful ingredient to construct inhomogeneous random graph sequences with nice statistical properties (edge independence, graph homomorphism, etc.). Most of the known literature on particle systems has been focused on the converging properties of particle systems on such random graph sequences, see, e.g., [1, 3, 17, 21]. They show that some of the classical mean-field arguments (as propagation of chaos [22]) can be extended to deal with random graphs and to include labeled graphons in the limit description. Nevertheless, only little attention has been put into the study of the limit object (1.4) which remains, to the author's knowledge, rather unknown.

The recent work [3] shows a direct connection between particle systems and the graphon theory: under suitable hypothesis on (1.2), there exists a Hölder-continuous mapping between the space of graphons $(\widetilde{\mathcal{W}}_0, \delta_\square)$ and $\tilde{\mu}$ as in Proposition 1.5. This allows to consider general graph sequences, as exchangeable random graphs [12], which leads to a (possibly) random graphon W in the limit. However, despite the rather understood convergence properties, only a few insights are available on the limit particle system, we refer to the examples in Subsection 2.3 of [3].

Although interesting on itself, the study of the limit Fokker-Planck equation does not necessarily provide a suitable description of particle systems on diverging time scales, as already raised in [9, 8] for the mean-field case. With the exception of dissipative dynamics [2], a substantial understanding of the phase space of (1.4) is needed to study the long-time behavior of finite particle systems on graphs as shown in [8]. It is thus important to better understand system (1.4) in order to address the finite particle system counterpart.

2. MAIN RESULT AND DISCUSSION

We suppose that the initial condition $\mu_0 \in \mathcal{P}(\mathbb{R} \times [0, 1])$ and the labeled graphon $W \in \mathcal{W}_0$ are fixed.

Before stating the main result, we introduce an equivalence relation on the unit interval and give the definition of degree in W . For x and y in $[0, 1]$, define the relation \sim by

$$x \sim y \quad \text{if and only if} \quad \mu_0(x) = \mu_0(y), \quad (2.1)$$

i.e., we identifies two labels whenever the initial conditions of the corresponding particles have the same law. Denote by J the quotient space $[0, 1]/\sim$ and, for $\bar{x} \in J$, denote its orbit by

$$[\bar{x}] := \{x \in [0, 1] : x \sim \bar{x}\}.$$

For some $x \in [0, 1]$, we denote the degree of x by $d(x) = \int_0^1 W(x, y)dy$. More generally, we define the degree of x with respect to a (measurable) subset $A \subset [0, 1]$ by

$$d_A(x) = \int_A W(x, y)dy. \quad (2.2)$$

Observe that under the measurability assumption (1.1), $[\bar{y}]$ is a measurable subset of $[0, 1]$ for every $\bar{y} \in J$; with a slight abuse of notation, we denote $d_{[\bar{y}]}(\cdot) := d_{[\bar{y}]}(\cdot)$.

2.1. Main result, examples and applications. We have the following theorem.

Theorem 2.1. *Assume Hypothesis 1.1. Suppose that, for every \bar{x} and \bar{y} in J it holds that*

$$d_{[\bar{y}]}(x) = \int_{[\bar{y}]} W(x, y)dy = \int_{[\bar{y}]} W(\bar{x}, y)dy = d_{[\bar{y}]}(\bar{x}) \quad \text{for all } x \in [\bar{x}]. \quad (2.3)$$

Then for every $\bar{x} \in J$, $\mu(x) = \mu(\bar{x})$, for all $x \in [\bar{x}]$.

Condition (2.3) is requiring that, if we partition the interval $[0, 1]$ in J subsets $\{[\bar{x}]\}_{\bar{x} \in J}$, then the density of neighbors with respect to any of these subsets, i.e., $d_{[\bar{y}]}(\cdot)$ for some $\bar{y} \in J$, is piecewise constant on $[0, 1]$ and completely characterized by the values on J . See Figure 1 (B) and Figure 2 (A) for graphons that satisfies (2.3).

Theorem 2.1 basically states that, if one can group the particles so that within each group they have the same initial condition (in law) and are *equi-connected* to the other groups in the sense of (2.3), then the solution μ to (1.4) is constant on these groups.

We observe that this result does not depend on the class of W as unlabeled graphon. As a consequence, one can relabel the particles such that μ is piecewise constant as a function of $x \in [0, 1]$.

Remark 2.2. *It is possible to weaken (2.3) by considering a larger J based on vertices which have not only the same initial law, but also the same degree in W . We refrain from increasing the complexity in the construction of J to keep the main ideas as clear as possible. This aspect is briefly discussed in the application to step kernels below.*

A direct consequence of Theorem 2.1 is that the information in $\{\mu(x)\}_{x \in [0, 1]}$ is contained in the possibly much smaller object $\{\mu(\bar{x})\}_{\bar{x} \in J}$.

Corollary 2.3. *Under the assumptions of Theorem 2.1. The infinite system (1.4) is suitably described by the (possible finite) system of coupled partial differential equations*

$$\begin{aligned} \partial_t \nu_t(\theta, \bar{x}) &= \frac{1}{2} \partial_\theta^2 (\sigma^2(\theta) \nu_t(\theta, \bar{x})) - \partial_\theta (\nu_t(\theta, \bar{x}) F(\theta)) \\ &\quad - \partial_\theta \left[\int_J W(\bar{x}, \bar{y}) \nu_t(\theta, \bar{x}) \int_{\mathbb{R}} \Gamma(\theta, \theta') \nu_t(d\theta', \bar{y}) d\bar{y} \right], \quad \bar{x} \in J, \end{aligned} \quad (2.4)$$

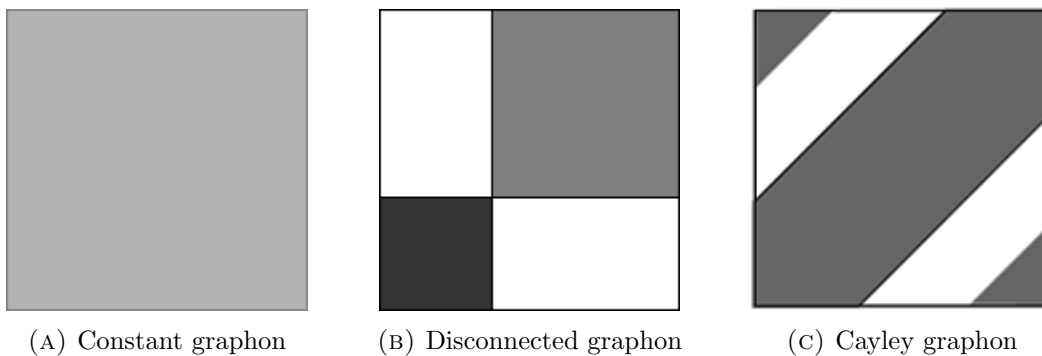


FIGURE 1. Examples of graphons that satisfy (2.5) (constant degree). Observe that the graphon in (B) is composed of two connected components, the smaller one being more densely connected (darker) with respect to the other one so to satisfy (2.5). Finally, observe that (A), (B) and (C) are different both as labeled and unlabeled graphons, i.e., different in δ_{\square} -distance.

with $\nu_0 = \mu_0$. Notably, $\mu(x) = \nu(\bar{x})$ for every $\bar{x} \in J$ and $x \in [\bar{x}]$.

Proof. Existence and uniqueness for system (2.4) directly follow from Proposition 1.3. Moreover, system (2.4) is written in closed form and can be thus solved independently of (1.4). \square

Graphons have proven to be an important tool for establishing the convergence of particle systems on graph sequences. Notably, it is now possible to show that interacting particle systems on a priori different graph sequences have the same asymptotic behavior whenever the limit of these sequences coincides as unlabeled graphon. However, different unlabeled graphons can lead to the same particle behavior: in these cases, the limit (1.4) tends to be a formal object instead of giving a precise description. Theorem 2.1 and Corollary 2.3 precise that both network structure and initial conditions are necessary to understand the system evolution.

We now turn to an interesting application when $|J| = 1$, i.e., when the law of the initial condition is label independent. It turns out that condition (2.3) boils down to a constant degree assumption on W .

Corollary 2.4. *Under the assumptions of Theorem 2.1. Suppose that $\mu_0(x) = \mu_0 \in \mathcal{P}(\mathbb{R})$ for every $x \in [0, 1]$. Then there exists $p \in [0, 1]$ such that condition (2.3) is equivalent to*

$$p = \int_0^1 W(x, y) dy, \quad \text{for all } x \in [0, 1]. \tag{2.5}$$

In particular, $\{\mu(x)\}_{x \in [0, 1]}$ is label independent, i.e., $\mu(\cdot) \equiv \mu \in \mathcal{P}(C([0, T], \mathbb{R}))$, and μ solves the classical McKean-Vlasov equation

$$\partial_t \nu_t(\theta) = \frac{1}{2} \partial_\theta^2 (\sigma^2(\theta) \nu_t(\theta)) - \partial_\theta (\nu_t(\theta) F(\theta)) - \partial_\theta \left[\nu_t(\theta) p \int_{\mathbb{R}} \Gamma(\theta, \theta') \nu_t(d\theta') \right], \tag{2.6}$$

with initial condition μ_0 .

Proof. The hypothesis on μ_0 forces $|J| = 1$ and thus condition (2.3) becomes (2.5). Corollary 2.3 yields the result. \square

The constant degree assumption on W is satisfied by a large class of non-trivial graphons as the examples shown in Figure 1.

Remark 2.5. *Observe that Corollary 2.4 can be derived from the results in [10] combined with the ones on particle systems on graphons, e.g., [17, 1]. Indeed, in [10] it is shown that, for a particle system defined on a graph sequence and with IID initial conditions, a sufficient condition to obtain the mean-field limit (2.6) is that each vertex in the (renormalized) graph sequence has the same asymptotic degree density. If one chooses such sequence to converge to a graphon with constant degree, then it satisfies both the hypothesis in [10, Theorem 1.1] and in, e.g., [1, Theorem 4.1]. Combining [10, Theorem 1.1] and [1, Theorem 4.1], we obtain that the same finite particle system converges to (2.6) but also to (1.4). As a consequence, the solutions of (1.4) and (2.6) must be the same, giving an undirect proof of Corollary 2.4.*

We step to a representative example whenever the graphon is a piecewise constant function, as the one in Figure 2.

Finite representation for step-kernels. For simplicity, we suppose that μ_0 is label independent, i.e., $\mu_0(x) = \mu_0 \in \mathcal{P}(\mathbb{R})$. However, observe that the strategy presented here can be applied by suitably approximating the initial condition with a constant-wise initial datum: indeed, the continuity of the solution μ to (1.4) with respect to the initial condition is a classical result [22].

Step kernels ([15, §7.1]) represent one of the fundamental blocks of graphon theory since they provide a way to approximate any $W \in \mathcal{W}_0$ with constant-wise functions. A function $W \in \mathcal{W}$ is a step kernel if there is a partition $\mathcal{P} = \{S_i\}_{i=1,\dots,k}$ of $[0, 1]$ into measurable sets such that W is constant on every product set $S_i \times S_j$. We use the following notation

$$W_{\mathcal{P}}(x, y) = \sum_{i,j=1}^k w_{ij} \mathbf{1}_{S_i \times S_j}(x, y), \quad \text{for } x, y \in [0, 1], \quad (2.7)$$

where $\{w_{ij}\}_{i,j=1,\dots,k}$ are bounded real numbers. See Figure 2 for an example with equidistant partition.

For a step kernel, condition (2.3) is clearly satisfied by taking $J = \{x_i\}_{i=1,\dots,k}$ with $x_i \in S_i$ for every $i = 1, \dots, k$, thus refining the strategy presented before Theorem 2.1, as anticipated in Remark 2.2.

Applying Theorem 2.1 and Corollary 2.3, we have the following representation for the graphon particle system (1.4) on the step graphon $W_{\mathcal{P}}$.

Proposition 2.6. *Suppose that $\mu_0 \in \mathcal{P}(\mathbb{R})$ and let W be the step kernel in (2.7). Then the graphon particle system (1.4) is suitably described by the finite collection of probability measures $\{\mu^i\}_{i=1,\dots,k}$ which solve*

$$\partial_t \mu_t^i(\theta) = \frac{1}{2} \partial_{\theta}^2 (\sigma^2(\theta) \mu_t^i(\theta)) - \partial_{\theta} (\mu_t^i(\theta) F(\theta)) - \partial_{\theta} \left[\sum_{j=1}^k w_{ij} \mu_t^i(\theta) \int_{\mathbb{R}} \Gamma(\theta, \theta') \mu_t^j(d\theta') \right], \quad (2.8)$$

and where $\mu_0^i = \mu_0$, for $i = 1, \dots, k$.

As shown in Proposition 2.6, the representation given in Corollary 2.3 becomes natural in the case of step kernels, and thus for graph limits arising from the stochastic block model. While Proposition 2.6 could be derived with direct computations², observe that

²A step kernel can be seen as a Stochastic Block Model, for which the representation (2.8) is somehow known.

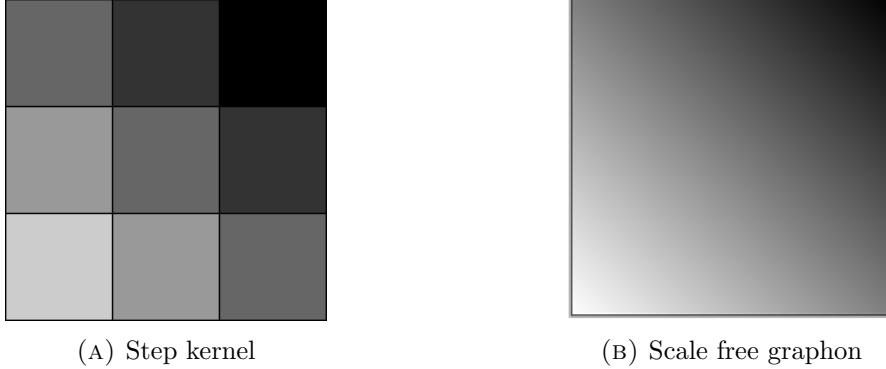


FIGURE 2. Suppose that $J = \{x_0, x_1, x_2\}$ and that $[x_0] = [0, 1/3)$, $[x_1] = [1/3, 2/3)$ and $[x_2] = [2/3, 1]$. Then, the step kernel in the figure (A) satisfies (2.3). It also represents a way to approximate the scale-free graphon (B). The grayscale stands for different values in $[0, 1]$.

Theorem 2.1 goes a step further: if one replaces every single constant block $S_i \times S_j$ in (2.7) by a (suitable scaled) graphon with constant degree, equation (2.8) does not change.

We also observe that combining Proposition 2.6 with the continuity estimates as in Remark 1.4, allows to approximate the graphon particle system (1.4) by a finite number of coupled Fokker-Planck equations. Since it is beyond the scope of this note, we do not pursue such analysis, yet we make this point a bit more precise in the next remark.

Remark 2.7. *By using the Weak Regularity Lemma [15, Corollary 9.13], one can approximate every graphon by a step kernel with an explicit control on the error. Assuming the continuity of (1.4) with respect to W , recall Remark 1.4, Proposition 2.6 opens the way to approximate the infinite graphon particle system (1.4) by using a finite (thus a priori numerically solvable) system of coupled Fokker-Planck equations, with precise bounds on the error.*

3. PROOF OF THE MAIN RESULT

3.1. Distance between probability measures. For two probability measures $\bar{\mu}, \bar{\nu} \in \mathcal{P}(C([0, T], \mathbb{R}))$, define their 2-Wasserstein distance as

$$D_T(\bar{\mu}, \bar{\nu}) = \inf_{X, Y} \left\{ \mathbf{E} \left[\sup_{t \in [0, T]} |X_t - Y_t|^2 \right] : \mathcal{L}(X) = \bar{\mu}, \mathcal{L}(Y) = \bar{\nu} \right\}^{1/2} \quad (3.1)$$

where the infimum is taken on all random variables X and Y with values in $C([0, T], \mathbb{R})$ and law \mathcal{L} equal to $\bar{\mu}$ and $\bar{\nu}$ respectively. From (3.1) we obtain that for every $s \in [0, T]$ and for every bounded 1-Lipschitz function f

$$\left| \int_{\mathbb{R}} f(\theta) \bar{\mu}_s(d\theta) - \int_{\mathbb{R}} f(\theta) \bar{\nu}_s(d\theta) \right| = \left| \int_{\mathbb{R}} f(\theta) [\bar{\mu}_s(d\theta) - \bar{\nu}_s(d\theta)] \right| \leq D_s(\bar{\mu}, \bar{\nu}). \quad (3.2)$$

Observe that (3.1) also makes sense with $T = 0$ and $C([0, T], \mathbb{R})$ replaced by \mathbb{R} .

3.2. Proof of Theorem 2.1. Recall the definition of the 2-Wasserstein distance D_T in (3.1), we aim at showing that

$$\max_{\substack{\bar{x} \in J \\ x \in [\bar{x}]} D_T(\mu(\bar{x}), \mu(x)) = 0. \quad (3.3)$$

We first assume $F \equiv 0$ and $\sigma \equiv 1$.

Let $\bar{x} \in J$ and $x \in [\bar{x}]$, without loss of generality, we can suppose that the associated realizations of the Brownian motion are the same. Using the fact that $x \in [\bar{x}]$, the initial conditions cancel and we have that

$$\begin{aligned} \theta_t^x - \theta_t^{\bar{x}} &= \int_0^t \int_0^1 W(x, y) \int_{\mathbb{R}} \Gamma(\theta_s^x, \theta) \mu(d\theta, y) dy ds \\ &\quad - \int_0^t \int_0^1 W(\bar{x}, y) \int_{\mathbb{R}} \Gamma(\theta_s^{\bar{x}}, \theta) \mu(d\theta, y) dy ds, \end{aligned} \quad (3.4)$$

In particular, this can be rewritten as

$$\begin{aligned} &\int_0^t \int_0^1 W(x, y) \int_{\mathbb{R}} [\Gamma(\theta_s^x, \theta) - \Gamma(\theta_s^{\bar{x}}, \theta)] \mu(d\theta, y) dy ds \\ &+ \int_0^t \int_0^1 [W(x, y) - W(\bar{x}, y)] \int_{\mathbb{R}} \Gamma(\theta_s^{\bar{x}}, \theta) \mu(d\theta, y) dy ds. \end{aligned} \quad (3.5)$$

We now use hypothesis (2.3) and add the following term in the previous equation

$$0 = \int_J \left[\int_{[\bar{y}]} [W(x, y) - W(\bar{x}, y)] \int_{\mathbb{R}} \Gamma(\theta_s^{\bar{x}}, \theta) \mu_s(d\theta, \bar{y}) dy \right] d\bar{y}, \quad (3.6)$$

where the integrals are sums whenever J or $[\bar{y}]$ are countable (recall that $[\bar{y}]$ is a measurable subset of $[0, 1]$ because of Hypothesis (1.1)).

By taking the squares and using $(a + b)^2 \leq 2(a^2 + b^2)$, this leads to

$$\begin{aligned} |\theta_t^x - \theta_t^{\bar{x}}|^2 &\leq 2T^2 \int_0^t |\theta_s^x - \theta_s^{\bar{x}}|^2 ds + \\ &+ 2T^2 \int_0^t \int_J \left| \int_{[\bar{y}]} [W(x, y) - W(\bar{x}, y)] \int_{\mathbb{R}} \Gamma(\theta_s^{\bar{x}}, \theta) [\mu_s(d\theta, y) - \mu_s(d\theta, \bar{y})] dy \right|^2 d\bar{y} ds, \end{aligned} \quad (3.7)$$

where we have used Cauchy-Schwartz inequality as well as the fact that Γ is 1-Lipschitz. Applying Cauchy-Schwartz again, and using (3.2), we are left with

$$|\theta_t^x - \theta_t^{\bar{x}}|^2 \leq 2T^2 \int_0^t |\theta_s^x - \theta_s^{\bar{x}}|^2 ds + 4T^2 \int_0^t \int_J \left[\int_{[\bar{y}]} D_s^2(\mu(y), \mu(\bar{y})) dy \right] d\bar{y} ds, \quad (3.8)$$

Thus, taking the supremum with respect to the time and the expected value \mathbf{E} this leads to

$$\begin{aligned} \mathbf{E} \left[\sup_{t \in [0, T]} |\theta_t^x - \theta_t^{\bar{x}}|^2 \right] &\leq 2T^2 \int_0^T \mathbf{E} \left[\sup_{u \in [0, s]} |\theta_u^x - \theta_u^{\bar{x}}|^2 \right] ds + \\ &+ 4T^2 \int_0^T \int_J \left[\int_{[\bar{y}]} D_s^2(\mu(y), \mu(\bar{y})) dy \right] d\bar{y} ds. \end{aligned} \quad (3.9)$$

Finally, we can take the maximum with respect to $\bar{x} \in J$ and $x \in [\bar{x}]$ and use the characterization of (3.1), to obtain

$$\max_{\substack{\bar{x} \in J \\ x \in [\bar{x}]}} D_T^2(\mu(\bar{x}), \mu(x)) \leq 6T^2 \int_0^T \max_{\substack{\bar{x} \in J \\ x \in [\bar{x}]}} D_s^2(\mu(\bar{x}), \mu(x)) ds, \quad (3.10)$$

which implies (3.3).

Whenever F is not zero, the proof is basically the same: using the Lipschitz properties of F , a term equal to $|\theta_s^x - \theta_s^{\bar{x}}|^2$ appears, thus adding the constant factor of $2T^2$ to the final equation (3.9).

If σ is not constant, then we may use the Burkholder-Davis-Gundy inequality (and the Lipschitz property of σ) to bound:

$$\mathbf{E} \left[\sup_{t \in [0, T]} \left| \int_0^t (\sigma(\theta_s^x) - \sigma(\theta_s^{\bar{x}})) dB_s \right|^2 \right] \leq CT^2 \int_0^T \mathbf{E} \left[\sup_{u \in [0, s]} |\theta_u^x - \theta_u^{\bar{x}}|^2 \right] ds,$$

with C a universal positive constant. The rest of the proof remains unchanged. \square

ACKNOWLEDGMENTS

The author is thankful to Gianmarco Bet, Giambattista Giacomini and Francesca Nardi for discussions while writing this note.

REFERENCES

- [1] E. Bayraktar, S. Chakraborty, and R. Wu. Graphon mean field systems. *arXiv:2003.13180 [math]*, Mar. 2020. arXiv: 2003.13180.
- [2] E. Bayraktar and R. Wu. Stationarity and uniform in time convergence for the graphon particle system. *arXiv:2008.10173 [math]*, 2020.
- [3] G. Bet, F. Coppini, and F. R. Nardi. Weakly interacting oscillators on dense random graphs. *arXiv:2006.07670 [math]*, June 2020. arXiv: 2006.07670.
- [4] S. Bhamidi, A. Budhiraja, and R. Wu. Weakly interacting particle systems on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 129(6):2174–2206, June 2019.
- [5] P. E. Caines and M. Huang. Graphon Mean Field Games and the GMFG Equations. *2018 IEEE Conference on Decision and Control (CDC)*, 2018.
- [6] R. Carmona, D. Cooney, C. Graves, and M. Lauriere. Stochastic Graphon Games: I. The Static Case. *arXiv:1911.10664 [math]*, Nov. 2019. arXiv: 1911.10664.
- [7] H. Chiba and G. S. Medvedev. The mean field analysis of the Kuramoto model on graphs i. The mean field equation and transition point formulas. *Discrete & Continuous Dynamical Systems - A*, 39(1):131, 2019.
- [8] F. Coppini. Long time dynamics for interacting oscillators on graphs. *arXiv:1908.01520 [math]*, Dec. 2019. arXiv: 1908.01520.
- [9] F. Coppini, H. Dietert, and G. Giacomini. A law of large numbers and large deviations for interacting diffusions on Erdős–Rényi graphs. *Stochastics and Dynamics*, 20(02):2050010, Apr. 2020.
- [10] S. Delattre, G. Giacomini, and E. Luçon. A Note on Dynamical Models on Random Graphs and Fokker-Planck Equations. *Journal of Statistical Physics*, 165(4):785–798, Nov. 2016.
- [11] J.-F. Delmas, D. Dronnier, and P.-A. Zitt. An Infinite-Dimensional SIS Model. *arXiv:2006.08241 [math]*, June 2020. arXiv: 2006.08241.
- [12] P. Diaconis and S. Janson. Graph limits and exchangeable random graphs. *Rendiconti di Matematica*, 28:33–61, 2008.
- [13] R. Dudley. *Real Analysis and Probability*. Cambridge University Press, 2 edition, 2002.
- [14] D. S. Kaliuzhnyi-Verbovetskyi and G. S. Medvedev. The Mean Field Equation for the Kuramoto Model on Graph Sequences with Non-Lipschitz Limit. *SIAM J. Math. Analysis*, 2017.
- [15] L. Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American mathematical society edition, 2012.

- [16] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, Nov. 2006.
- [17] E. Luçon. Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 2020.
- [18] E. Luçon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, 24(5):1946–1993, Oct. 2014.
- [19] G. S. Medvedev. The continuum limit of the Kuramoto model on sparse random graphs. *arXiv:1802.03787 [nlin]*, Feb. 2018. arXiv: 1802.03787.
- [20] F. Parise and A. Ozdaglar. Graphon games: A statistical framework for network games and interventions. *arXiv:1802.00080 [cs]*, June 2020. arXiv: 1802.00080.
- [21] G. H. Reis and R. I. Oliveira. Interacting diffusions on random graphs with diverging degrees: hydrodynamics and large deviations. *Journal of Statistical Physics*, 2019. arXiv: 1807.06898.
- [22] A.-S. Sznitman. Topics in propagation of chaos. In P.-L. Hennequin, editor, *Ecole d'Été de Probabilités de Saint-Flour XIX - 1989*, volume 1464, pages 165–251. Springer Berlin Heidelberg, 1991.