MEAN-FIELD THEORY VIA DISSOCIATED ARRAYS FOR PARTICLE SYSTEMS INTERACTING THROUGH NOISY WEIGHTS

NICOLAS FOURNIER AND DATONG ZHOU

Abstract. We study a mean-field limit for a N-particle system in which each particle follows a diffusion and interacts with other particles through a weight on each directed edge. Each weight evolves according to its own nonlinear SDE driven by a Brownian motion, with coefficients involving the states of the two endpoint particles of the edge. The initial vertex and edge variables are assumed to have a dissociated Aldous–Hoover form. We construct the limiting nonlinear SDE by averaging the interaction over an independent neighbor and an edge input, prove its well-posedness, and show that the dissociated vertex-edge structure is propagated by the dynamics. This propagation property is an analogue of propagation of chaos in the case where the weight of each edge may remain correlated with the states of the two endpoint particles. Under either a bounded-observable assumption or a sub-Gaussian edge-input condition, the finite system converges to this limit through quantitative coupling estimates for a typical particle and a typical edge. We also prove the convergence of the empirical measure of particle's state pairs and their interaction weights.

1. Description of the model and main result

1.1. **Motivation.** In its classical exchangeable form, mean-field theory describes the large-population limit of an interacting particle system through one typical particle, of which the state solves a nonlinear SDE, see the seminal works of Kac [22], McKean [27], Sznitman [30], Méléard [28].

A first way to incorporate heterogeneity is to keep the interaction structure fixed: the interaction kernel is prescribed in advance. Graphon and graph-kernel mean-field limits describe such system, see for instance Chiba and Medvedev [9], Kaliuzhnyi-Verbovetskyi and Medvedev [23], Bayraktar, Chakraborty and Wu [3] and Jabin and co-authors [18, 20, 19]. In all these works, the interaction weights remain fixed along the dynamics.

However, in many adaptive network models, the interaction strength between two particles is itself a dynamical variable, depending on the states of the two particles and possibly driven by its own noise. Such feedback is natural in co-evolutionary network models and in synaptic plasticity, see Gross and Blasius [16] and Bi and Poo [5].

Mathematical works on genuine co-evolutionary dynamics include deterministic continuum limits for pairwise adaptive networks, see Gkogkas, Kuehn and Xu [14, 15], as well as probabilistic models with intrinsic stochasticity and dynamically evolving edge variables, see Bayraktar and Wu [4] and Ganguly, Spiliopoulos and Sussman [13]. The companion work [31] develops a structural metric and sampling viewpoint for linear weight adaptation.

The goal of the present paper is to treat a class of particle systems with fully noisy and nonlinear real-valued edge weights. We keep track not only of a one-particle law, but of a vertex-edge array whose dissociated structure has to be propagated by the dynamics.

2020 *Mathematics Subject Classification.* 60K35, 60H10, 49Q22.

Key words and phrases. Interacting particle systems, Dissociated arrays, Mean-field limits, McKean-Vlasov equations.

1.2. **The interacting particle system.** We fix $d \geq 1$ and $m \geq 1$. For each $N \geq 2$, we set $I_N = \{1, \dots, N\}$ and consider the system

$$X_t^{i,N} = X_0^i + \int_0^t b(X_s^{i,N}) \mathrm{d}s + \int_0^t \sigma(X_s^{i,N}) \mathrm{d}B_s^i + \frac{1}{N-1} \sum_{j \in I_N, j \neq i} \int_0^t \phi(U_s^{ij,N}) \gamma(X_s^{i,N}, X_s^{j,N}) \mathrm{d}s, \tag{1}$$

$$U_t^{ij,N} = U_0^{ij} + \int_0^t \alpha(X_s^{i,N}, X_s^{j,N}, U_s^{ij,N}) \mathrm{d}s + \int_0^t \beta(X_s^{i,N}, X_s^{j,N}, U_s^{ij,N}) \mathrm{d}W_s^{ij}. \tag{2}$$

The first equation (1) has to hold for all $i \in I_N$, the second one (2) for all $i, j \in I_N$ with $i \neq j$.

1.3. **Hypotheses.** Concerning randomness, we suppose the following.

Assumption 1. *The following random objects are independent:*

- *an i.i.d. family $(\xi_i)_{i \geq 1}$ with common law π valued in some measurable space $(E, \mathcal{E})$,*
- *an i.i.d. family $(\xi_{ij}^\dagger)_{i,j \geq 1, i \neq j}$ with common law $\pi^\dagger$ valued in some measurable space $(E^\dagger, \mathcal{E}^\dagger)$,*
- *some i.i.d. m-dimensional Brownian motions $(B_t^i)_{t \geq 0, i \geq 1}$,*
- *some i.i.d. 1-dimensional Brownian motions $(W_t^{ij})_{t \geq 0, i,j \geq 1, i \neq j}$.*

There are some measurable functions $F : E \to \mathbb{R}^d$ and $G : E^2 \times E^\dagger \to \mathbb{R}$ such that, for all $i \geq 1$, $X_0^i = F(\xi_i)$ and, for all $i, j \geq 1$ with $i \neq j$, $U_0^{ij} = G(\xi_i, \xi_j, \xi_{ij}^\dagger)$.

This is the standard dissociated Aldous–Hoover form for the initial edge weights: each edge ij has its own mark $\xi_{ij}^\dagger$, independent of the vertex variables ξ_i and ξ_j and of the other edge marks. For the representation-theoretic background, see Aldous [1], Hoover [17] and Kallenberg[24, 25]; for the connection with graph limits, see Janson and Diaconis [21]. In the present nonlinear setting, the edge mark is not merely a technical decoration: it records edge-level randomness which can be seen by nonlinear functions of the weight and which later evolves together with the edge Brownian motion.

We will assume at least the following moment condition.

Assumption 2. *For $\xi_1, \xi_2, \xi_{12}^\dagger$ independent with $\xi_1, \xi_2 \sim \pi$ and $\xi_{12}^\dagger \sim \pi^\dagger$, it holds that*

$$\mathbb{E}[|F(\xi_1)|^2 + (G(\xi_1, \xi_2, \xi_{12}^\dagger))^2] < \infty.$$

Concerning the coefficients, we assume the following.

Assumption 3. *The functions $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathcal{M}_{d \times m}(\mathbb{R})$, $\gamma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $\alpha : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ and $\beta : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ are bounded and globally Lipschitz continuous, while $\phi : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous.*

We keep ϕ merely Lipschitz at this stage, in order to include the natural choice $\phi(u) = u$. Additional boundedness or moment assumptions are stated explicitly in the results below.

Remark 4. *Under Assumptions 1, 2 and 3, the system (1)-(2) has a pathwise unique strong solution, because it can be written as an SDE in $\mathbb{R}^{Nd+N(N-1)}$ with locally Lipschitz continuous coefficients with at most linear growth.*

1.4. **The mean-field limit.** We aim to study what happens as N tends to infinity. The mean-field limit is described by a nonlinear SDE coupled with an averaged edge interaction. To rigorously define a limit process, we need the following result, that will be proved in Section 3 using some arguments found in Karandikar [26]. It provides a measurable solution map for the edge equation.

Lemma 5. *Grant Assumptions 1, 2 and 3 and set $\mathcal{C}_\ell = C(\mathbb{R}_+, \mathbb{R}^\ell)$. There is a measurable map*

$$\Gamma = (\Gamma_t)_{t \geq 0} : E \times \mathcal{C}_d \times E \times \mathcal{C}_d \times E^\dagger \times \mathcal{C}_1 \to \mathcal{C}_1$$

such that for any space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, any triple $(\xi, \xi', \xi^\dagger)$ of independent $\mathcal{F}_0$-measurable variables with $\xi, \xi' \sim \pi$ and $\xi^\dagger \sim \pi^\dagger$, any pair of continuous $(\mathcal{F}_t)_{t\geq 0}$-adapted $\mathbb{R}^d$-valued processes $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ such that

$$\forall\, T > 0,\ \exists\, A_T > 0 \text{ such that } \forall\, 0 \leq s < t < s+1 \leq T+1, \quad \mathbb{E}[|X_t - X_s|^2] \leq A_T(t-s), \tag{3}$$

and such that Y satisfies the same condition, any 1-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $W = (W_t)_{t\geq 0}$, the pathwise unique solution $U = (U_t)_{t\geq 0}$ to

$$U_t = G(\xi, \xi', \xi^\dagger) + \int_0^t \alpha(X_s, Y_s, U_s)\mathrm{d}s + \int_0^t \beta(X_s, Y_s, U_s)\mathrm{d}W_s \tag{4}$$

satisfies $U = \Gamma(\xi, X, \xi', Y, \xi^\dagger, W)$ in the sense that a.s., for all $t \geq 0$, $U_t = \Gamma_t(\xi, X, \xi', Y, \xi^\dagger, W)$. One may moreover build Γ in such a way that for all random sextuplet $(\xi, X, \xi', Y, \xi^\dagger, W)$ as above, a.s.,

$$\text{for all } t \geq 0, \quad \Gamma_t(\xi, X, \xi', Y, \xi^\dagger, W) = \Gamma_t(\xi, (X_{s\wedge t})_{s\geq 0}, \xi', Y, \xi^\dagger, W). \tag{5}$$

We call a pair (ξ, X) admissible if it is defined on some filtered probability space, if $\xi \sim \pi$ is $\mathcal{F}_0$-measurable, and if $X = (X_t)_{t\geq 0}$ is continuous, adapted and satisfies (3). A probability measure μ on $E \times \mathcal{C}_d$ is called admissible if it is the law of an admissible pair.

The next lemma, to be proved in Section 3, introduces the averaged edge-interaction operator. Informally, this operator takes the edge contribution seen from an admissible input (ξ, X), computed through the map Γ from Lemma 5 and the nonlinearity ϕ, and averages it over an independent neighbor with law μ and over the fixed edge input law $\pi^\dagger \otimes \mathbb{W}$, where $\mathbb{W}$ is the Wiener measure on $\mathcal{C}_1$.

Lemma 6. *Grant Assumptions 1, 2 and 3 and suppose either that ϕ is bounded or that*

$$\sup_{a,b\in E} \int_{E^\dagger} |G(a,b,c)|^2 \pi^\dagger(\mathrm{d}c) < \infty. \tag{6}$$

For any admissible probability measure μ on $E \times \mathcal{C}_d$, we use the pointwise notation, for $t \geq 0$ and $(a, x) \in E \times \mathcal{C}_d$,

$$\Lambda_\mu(t, a, x) = \int_{E\times\mathcal{C}_d} \int_{E^\dagger\times\mathcal{C}_1} \phi(\Gamma_t(a, x, b, y, c, w))\gamma(x_t, y_t)\pi^\dagger(\mathrm{d}c)\mathbb{W}(\mathrm{d}w)\mu(\mathrm{d}b, \mathrm{d}y).$$

(i) For each admissible pair (ξ, X), the map $t \mapsto \Lambda_\mu(t, \xi, X)$ is a.s. continuous.

(ii) If (ξ, X) is admissible with respect to $(\mathcal{F}_t)_{t\geq 0}$, then $(\Lambda_\mu(t, \xi, X))_{t\geq 0}$ is $(\mathcal{F}_t)_{t\geq 0}$-adapted.

For (ξ, X) an admissible pair, we use the shorthand

$$\Lambda_{\xi,X}(t, a, x) := \Lambda_{\mathrm{Law}(\xi,X)}(t, a, x).$$

In particular, $\Lambda_{\xi,X}(t, \xi, X)$ means that the law in the operator is $\mathrm{Law}(\xi, X)$ and that the operator is evaluated at (ξ, X). We now introduce the limiting nonlinear SDE corresponding to (1)-(2).

Definition 7. *Grant Assumptions 1, 2 and 3 and suppose either that ϕ is bounded or that (6) holds. We say that (ξ, X) defined on some $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ solves the nonlinear SDE if $\xi \sim \pi$ is $\mathcal{F}_0$-measurable, if $X = (X_t)_{t\geq 0}$ is continuous, $\mathbb{R}^d$-valued, $(\mathcal{F}_t)_{t\geq 0}$-adapted and satisfies (3), and if there is a m-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $B = (B_t)_{t\geq 0}$ such that a.s., for all $t \geq 0$,*

$$X_t = F(\xi) + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s + \int_0^t \Lambda_{\xi,X}(s, \xi, X)\mathrm{d}s. \tag{7}$$

In words, X is the limiting trajectory of a typical particle. Its dynamics contains the local drift b, the Brownian noise σdB, and the averaged interaction field $\Lambda_{\xi,X}$, which is obtained by averaging over an independent typical neighbor, together with an additional independent edge variable and an independent edge Brownian motion.

1.5. **Results.** Our first main result establishes well-posedness for the limiting equation before any finite-N comparison is made. The averaged operator $\Lambda_{\xi,X}$ depends on the law of the candidate process itself, so the statement is a McKean–Vlasov well-posedness result for the vertex component.

Theorem 8. *Grant Assumptions 1, 2 and 3 and suppose that either ϕ is bounded or (6).*

(i) For any stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, any $\mathcal{F}_0$-measurable random variable ξ with law π, and any m-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion B, there exists a pathwise unique continuous adapted process X such that (ξ, X) solves the nonlinear SDE driven by B.

(ii) Uniqueness in law holds: if (ξ, X) and $(\tilde{\xi}, \tilde{X})$ are two solutions, possibly defined on different stochastic bases, then $\mathrm{Law}(\xi, X) = \mathrm{Law}(\tilde{\xi}, \tilde{X})$.

Our second result concerns the limit $N \to \infty$. The comparison process is obtained by solving the nonlinear SDE for each vertex input (ξ_i, B^i) and then, for each directed edge, solving the edge equation driven by the corresponding edge input $(\xi^\dagger_{ij}, W^{ij})$. Thus the limiting object is an explicit dissociated vertex-edge array, built on the same sources of randomness as the particle system. The theorem first gives pathwise coupling estimates for a typical particle and a typical edge, and then projects this coupling to the empirical law of triples $(X^{i,N}_t, X^{j,N}_t, U^{ij,N}_t)$.

Theorem 9. *Grant Assumptions 1, 2 and 3 and suppose that either ϕ is bounded or (6) and*

$$\text{there is } \theta_0 > 0 \text{ such that } \int_{E^2\times E^\dagger} \exp\big(\theta_0 |G(a,b,c)|^2\big)\pi(\mathrm{d}a)\pi(\mathrm{d}b)\pi^\dagger(\mathrm{d}c) < \infty. \tag{8}$$

For each $i \geq 1$, consider the solution $\bar{X}^i = (\bar{X}^i_t)_{t\geq 0}$ to the nonlinear SDE corresponding to ξ_i and B^i. For $i, j \geq 1$ with $i \neq j$, consider the pathwise unique solution $\bar{U}^{ij} = (\bar{U}^{ij}_t)_{t\geq 0}$ to

$$\bar{U}^{ij}_t = G(\xi_i, \xi_j, \xi^\dagger_{ij}) + \int_0^t \alpha(\bar{X}^i_s, \bar{X}^j_s, \bar{U}^{ij}_s)\mathrm{d}s + \int_0^t \beta(\bar{X}^i_s, \bar{X}^j_s, \bar{U}^{ij}_s)\mathrm{d}W^{ij}_s. \tag{9}$$

(i) If ϕ is bounded, then for all $T > 0$, there is a constant $C_T > 0$ such that for all $N \geq 2$,

$$\mathbb{E}\Big[\sup_{t\in[0,T]} (|X^{1,N}_t - \bar{X}^1_t|^2 + |U^{12,N}_t - \bar{U}^{12}_t|^2)\Big] \leq \frac{C_T}{N}.$$

(ii) If (6) and (8) hold, then for all $T > 0$, there is $C_T > 0$ such that for all $N \geq 2$,

$$\mathbb{E}\Big[\sup_{t\in[0,T]} (|X^{1,N}_t - \bar{X}^1_t| + |U^{12,N}_t - \bar{U}^{12}_t|)\Big] \leq \frac{C_T \exp(C_T\sqrt{\log N})}{N^{\frac{1}{2}}}.$$

(iii) In either case above, introduce $f_t = \mathrm{Law}(\bar{X}^1_t, \bar{X}^2_t, \bar{U}^{12}_t)$. For all $T > 0$, there is $C_T > 0$ such that for all $N \geq 3$,

$$\sup_{t\in[0,T]} \mathbb{E}\Big[\mathcal{W}_1\Big(\frac{1}{N(N-1)} \sum_{i,j\in I_N, i\neq j} \delta_{(X^{i,N}_t, X^{j,N}_t, U^{ij,N}_t)}, f_t\Big)\Big] \leq \frac{C_T}{N^{\frac{1}{2d+1}}}.$$

The Wasserstein distance $\mathcal{W}_1$ between two probability measures f, g on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$ is defined by

$$\mathcal{W}_1(f,g) = \inf\Big\{\mathbb{E}[|X - \tilde{X}| + |Y - \tilde{Y}| + |U - \tilde{U}|] : (X,Y,U) \sim f, (\tilde{X}, \tilde{Y}, \tilde{U}) \sim g\Big\}.$$

The first propagation estimate is in L^2 and uses the boundedness of ϕ. When ϕ is unbounded but Lipschitz, the proof instead uses a sub-Gaussian control of the edge variable and yields the weaker L^1 rate in (ii). The same truncation strategy could be adapted to an L^2 estimate, but this would produce a much worse rate than in the L^1 estimate above. We therefore state only the L^1 bound. Point (iii) then translates these couplings into the edge time-marginals analogue of classical propagation of chaos.

1.6. **Propagation of dissociatedness.** In the classical setting, propagation of chaos is represented by independent copies of a single nonlinear process. Here the limiting object is an array: subarrays supported on disjoint vertex sets are independent, but an edge remains coupled with its two endpoints. The next remark makes explicit how this Aldous–Hoover structure is propagated by the limiting dynamics.

Remark 10. *After fixing measurable versions of the strong solution maps, we may write*

$$\bar{X}^i = \mathsf{F}(\xi_i, B^i), \qquad \bar{U}^{ij} = \mathsf{G}(\xi_i, B^i, \xi_j, B^j, \xi^\dagger_{ij}, W^{ij}),$$

where $\mathsf{F} : E \times \mathcal{C}_m \to \mathcal{C}_d$ *is the solution map for the nonlinear SDE and*

$$\mathsf{G}(a, v, b, \tilde{v}, c, w) = \Gamma(a, \mathsf{F}(a, v), b, \mathsf{F}(b, \tilde{v}), c, w).$$

The dissociatedness (Aldous–Hoover representation) of Assumption 1 is therefore propagated by the limit dynamics: vertices carry (ξ_i, B^i) *and directed edges carry* $(\xi^\dagger_{ij}, W^{ij})$*. Hence, for any* $t \geq 0$*, for any collection* $C \subset \{(i,j) : i \geq 1, j \geq 1, i \neq j\}$ *such that for all* $(i,j),(i',j') \in C$ *distinct, the four numbers* i, j, i', j' *are distinct, the variables* $(\bar{X}^i_t, \bar{X}^j_t, \bar{U}^{ij}_t)_{(i,j)\in C}$ *are i.i.d. This propagation of dissociatedness is the analogue of the classical propagation of chaos.*

1.7. **Comments.** One line of non-exchangeable mean-field theory starts from a fixed graph kernel or graphon and derives an extended Vlasov or McKean–Vlasov equation, see for instance the works cited above [9, 23, 3, 18, 20]. This is particularly well-suited to non-adaptive interaction structures, with fixed interaction weights. Here the fixed-label picture means the following type of representation: one embeds the vertices into a predetermined latent space D and represents the edge structure at time t by a deterministic kernel $w_t : D \times D \to \mathbb{R}$, evaluated at the two labels. The label of a vertex is chosen once and for all, and the random evolution is not allowed to enlarge the domain by adding new vertex or edge randomness.

A second line studies dynamic random networks using exchangeable arrays and graph limits, closely connected with the Aldous–Hoover viewpoint; see Crane [10, 11], Černỳ and Klimovsky [8], Athreya, Den Hollander and Röllin [2] and Braunsteins, den Hollander and Mandjes [7, 6]. These papers actually study some graph-valued Markov processes. There are no underlying particles.

Closer to our setting are works on genuine co-evolutionary dynamics, where the edge variables and the vertex states influence each other. In the deterministic case, Gkogkas, Kuehn and Xu [14, 15] establish continuum and mean-field limits for adaptive Kuramoto-type networks through evolving kernels, graph measures, or signed digraph measures. These models share the pairwise feedback mechanism with ours, but they are fully deterministic systems.

Probabilistic models with intrinsic stochasticity have also been considered. Bayraktar and Wu [4] study a continuous-time model in which both the vertex and edge states take values in some countable spaces, while Ganguly, Spiliopoulos and Sussman [13] analyze a discrete-time latent-variable model with binary dynamic edges, feedback effects, and graphon/multiplexon limits.

The companion work [31] develops a structural metric and sampling viewpoint for the same system (1)-(2) when the weight adaptation is linear and non noisy (that is when $\phi(u) = u$, $\alpha(x,y,u) = a(x,y) + b(x,y)u$ and $\beta(x,y,u) = 0$). In that setting, weight fluctuations can be absorbed at the level of the empirical state-weight structure. The main stability estimate in [31] controls the evolution in a hybrid Wasserstein and cut distance by the initial structural error and the graphon sampling error.

Here we allow a more general edge dynamics, where the averaging reduction is no longer available. The edge variables solve their own noisy dynamics and enter the particle equation nonlinearly, so the full edge law and its correlation with the endpoint trajectories have to be retained in the mean-field description. To our knowledge, this is the first mean-field result for fully nonlinear and noisy adaptive weights.

1.8. **Technical comments.** We finish with a few comments on possible extensions and on the rates obtained above. First, it would be natural to relax the boundedness assumptions on the coefficients, while keeping global Lipschitz continuity. This should be possible under suitable additional moment assumptions, but would require more involved estimates.

The convergence rate in Theorem 9-(iii) is not intended to be optimal. Our proof partitions the directed edges, so that each independent subsample has only order N (independent) elements. Applying the standard empirical $\mathcal{W}_1$ estimate in dimension $2d+1$, see [12, Theorem 1], gives the rate $N^{-1/(2d+1)}$. In the idealized situation where the $N(N-1)$ edge triples were independent, the same general empirical $\mathcal{W}_1$ estimate would give a rate of order $N^{-2/(2d+1)}$. We do not know what the optimal rate should be.

A related extension would be to replace the diffusive particle dynamics by jump-type dynamics, as in integrate-and-fire models from mathematical neuroscience. Such models are often driven by Poisson events rather than Brownian noises, see for instance [19, 20]. We expect that the same dissociated vertex-edge viewpoint should remain useful.

1.9. **Plan of the paper.** Section 2 collects the estimates needed for the nonlinear operator Λ and then proves the well-posedness and propagation results. Section 3 proves the measurability and non-anticipativity statements for the edge solution map Γ and the averaged operator Λ.

2. Main proofs

2.1. **Estimates.** The purpose of this subsection is to keep the inequality estimates separate from the fixed-point and propagation arguments. We use the map Γ from Lemma 5 and the definition of Λ_μ; the estimates collected here will then be used in the proof of Lemma 6. We first record estimates for the edge map Γ, with one auxiliary space for the edge mark $\xi^\dagger$ and another one for the Brownian input W. We then derive the corresponding estimates for the averaged operator Λ_μ.

Proposition 11. *Grant Assumptions 1, 2 and 3. For all $T>0$, there is $C_T>0$, depending only on T,α,β, such that the following estimates hold.*

(i) For every pair of independent admissible inputs (ξ,X) and (ξ',Y) defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, every auxiliary edge-mark probability space $(\Omega^\dagger,\mathcal{F}^\dagger,\mathbb{P}^\dagger)$ carrying an $E^\dagger$-valued random variable $\xi^\dagger$ with law $\pi^\dagger$, and every auxiliary Brownian space $(\Omega^\#,\mathcal{F}^\#,(\mathcal{F}^\#_t)_{t\ge0},\mathbb{P}^\#)$ carrying a one-dimensional Brownian motion W, all independent of each other,

$$\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\sup_{t\in[0,T]}|\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W)|^2\Big]\Big]\le 2\mathbb{E}^\dagger\big[|G(\xi,\xi',\xi^\dagger)|^2\big]+C_T.$$

(ii) For every pair of vertex-side triples $(\xi,X,\tilde X)$ and $(\xi',Y,\tilde Y)$ such that (ξ,X), $(\xi,\tilde X)$, (ξ',Y) and $(\xi',\tilde Y)$ are admissible and such that $(\xi,X,\tilde X)$ is independent of $(\xi',Y,\tilde Y)$, and every pair of auxiliary spaces as in (i), independent of these triples, for all $t\in[0,T]$,

$$\begin{aligned}&\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\sup_{u\in[0,t]}|\Gamma_u(\xi,X,\xi',Y,\xi^\dagger,W)-\Gamma_u(\xi,\tilde X,\xi',\tilde Y,\xi^\dagger,W)|^2\Big]\Big]\\ \le& C_T\Big(\sup_{s\in[0,t]}|X_s-\tilde X_s|^2+\sup_{s\in[0,t]}|Y_s-\tilde Y_s|^2\Big).\end{aligned}$$

Proof. For (i), set $U_t=\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W)$, $t\ge0$. Lemma 5 gives

$$U_t=G(\xi,\xi',\xi^\dagger)+\int_0^t\alpha(X_s,Y_s,U_s)\mathrm{d}s+\int_0^t\beta(X_s,Y_s,U_s)\mathrm{d}W_s.$$

Since α and β are bounded, the Cauchy–Schwarz inequality for the drift term, and the Burkholder-Davies-Gundy inequality for the martingale term, give

$$\mathbb{E}^\#\Big[\sup_{t\in[0,T]}|U_t-G(\xi,\xi',\xi^\dagger)|^2\Big]\le C_T.$$

Hence

$$\mathbb{E}^{\#}\Big[\sup_{t\in[0,T]}|U_t|^2\Big] \le 2|G(\xi,\xi',\xi^\dagger)|^2 + C_T,$$

and applying $\mathbb{E}^\dagger$ allows us to conclude.

For (ii), set

$$U_r = \Gamma_r(\xi, X, \xi', Y, \xi^\dagger, W), \qquad \tilde{U}_r = \Gamma_r(\xi, \tilde{X}, \xi', \tilde{Y}, \xi^\dagger, W), \qquad r \ge 0.$$

Recalling Lemma 5, we see that

$$U_r - \tilde{U}_r = \int_0^r [\alpha(X_s, Y_s, U_s) - \alpha(\tilde{X}_s, \tilde{Y}_s, \tilde{U}_s)]\mathrm{d}s + \int_0^r [\beta(X_s, Y_s, U_s) - \beta(\tilde{X}_s, \tilde{Y}_s, \tilde{U}_s)]\mathrm{d}W_s.$$

Using the Lipschitz continuity of α, β, the Cauchy–Schwarz inequality for the drift term, and the Burkholder-Davies-Gundy inequality for the martingale term, we get, for $r \le T$,

$$\mathbb{E}^{\#}\Big[\sup_{u\in[0,r]}|U_u - \tilde{U}_u|^2\Big] \le C_T \int_0^r \Big(|X_s - \tilde{X}_s|^2 + |Y_s - \tilde{Y}_s|^2 + \mathbb{E}^{\#}\Big[\sup_{u\in[0,s]}|U_u - \tilde{U}_u|^2\Big]\Big)\mathrm{d}s.$$

Thanks to Gronwall's lemma,

$$\mathbb{E}^{\#}\Big[\sup_{u\in[0,r]}|U_u - \tilde{U}_u|^2\Big] \le C_T\Big(\sup_{s\in[0,t]}|X_s - \tilde{X}_s|^2 + \sup_{s\in[0,t]}|Y_s - \tilde{Y}_s|^2\Big).$$

Applying $\mathbb{E}^\dagger$ allows us to conclude. □

Proposition 12. *Grant Assumptions 1, 2 and 3 and suppose either that ϕ is bounded or that (6) holds. For all $T > 0$, there is $C_T > 0$ such that the following estimates hold.*

(i) For any admissible probability measure μ on $E \times \mathcal{C}_d$, any admissible pair (ξ, X), any $t \in [0,T]$,

$$|\Lambda_\mu(t, \xi, X)| \le C_T \quad a.s.$$

(ii) Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ support an $\mathcal{F}_0$-measurable random variable ξ and continuous adapted processes $X = (X_t)_{t\ge0}$ and $\tilde{X} = (\tilde{X}_t)_{t\ge0}$ such that (ξ, X) and $(\xi, \tilde{X})$ are admissible. Write $\Lambda_{\xi,X} := \Lambda_{\mathrm{Law}(\xi,X)}$ and $\Lambda_{\xi,\tilde{X}} := \Lambda_{\mathrm{Law}(\xi,\tilde{X})}$. Then, for all $t \in [0,T]$,

$$\mathbb{E}\Big[|\Lambda_{\xi,X}(t,\xi,X) - \Lambda_{\xi,\tilde{X}}(t,\xi,\tilde{X})|^2\Big] \le C_T \mathbb{E}\Big[\sup_{s\in[0,t]}|X_s - \tilde{X}_s|^2\Big].$$

Proof. We start with (i). Additionally to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ on which is defined (ξ, X), we introduce a probability space $(\Omega^*, \mathcal{F}^*, (\mathcal{F}^*_t)_{t\ge0}, \mathbb{P}^*)$ endowed with an admissible pair (ξ', Y) with law μ, a probability space $(\Omega^\dagger, \mathcal{F}^\dagger, \mathbb{P}^\dagger)$ endowed with an $E^\dagger$-valued random variable $\xi^\dagger$ with law $\pi^\dagger$, and a probability space $(\Omega^{\#}, \mathcal{F}^{\#}, (\mathcal{F}^{\#}_t)_{t\ge0}, \mathbb{P}^{\#})$ endowed with a one-dimensional Brownian motion W. The three auxiliary spaces are taken independent of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$ and of each other. Recalling Lemma 6,

$$\Lambda_\mu(t, \xi, X) = \mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^{\#}\Big[\phi(\Gamma_t(\xi, X, \xi', Y, \xi^\dagger, W))\gamma(X_t, Y_t)\Big]\Big]\Big].$$

Define

$$\Psi_{\mu,T}(\xi, X) := \mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^{\#}\Big[\sup_{t\in[0,T]}|\phi(\Gamma_t(\xi, X, \xi', Y, \xi^\dagger, W))|\Big]\Big]\Big].$$

If ϕ is bounded, then $\Psi_{\mu,T}(\xi, X) \le ||\phi||_\infty$. If (6) holds, then we use that $|\phi(u)| \le C(1 + |u|)$, Proposition 11-(i), together with the Cauchy–Schwarz inequality. This gives

$$\Psi_{\mu,T}(\xi, X) \le C + C\mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^{\#}\Big[\sup_{t\in[0,T]}|\Gamma_t(\xi, X, \xi', Y, \xi^\dagger, W)|\Big]\Big]\Big] \le C + C_T \mathbb{E}^*[\mathbb{E}^\dagger[|G(\xi,\xi',\xi^\dagger)|^2]]^{\frac{1}{2}}.$$

This last quantity is bounded by some constant C_T thanks to (6).

Then (i) follows from $|\Lambda_\mu(t, \xi, X)| \le ||\gamma||_\infty \Psi_{\mu,T}(\xi, X)$.

We now prove (ii). Additionally to $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$, we introduce a second probability space $(\Omega^*,\mathcal{F}^*,(\mathcal{F}^*_t)_{t\geq 0},\mathbb{P}^*)$ endowed with a copy $(\xi^*,X^*,\tilde X^*)$ of $(\xi,X,\tilde X)$, as well as a third probability space $(\Omega^\dagger,\mathcal{F}^\dagger,\mathbb{P}^\dagger)$ endowed with an $E^\dagger$-valued random variable $\xi^\dagger$ with law $\pi^\dagger$ and a fourth probability space $(\Omega^\#,\mathcal{F}^\#,(\mathcal{F}^\#_t)_{t\geq 0},\mathbb{P}^\#)$ endowed with a one-dimensional Brownian motion W. On

$$(\bar\Omega,\bar{\mathcal{F}},(\bar{\mathcal{F}}_t)_{t\geq 0},\bar{\mathbb{P}})=\big(\Omega\times\Omega^*\times\Omega^\dagger\times\Omega^\#,\mathcal{F}\otimes\mathcal{F}^*\otimes\mathcal{F}^\dagger\otimes\mathcal{F}^\#,(\mathcal{F}_t\otimes\mathcal{F}^*_t\otimes\mathcal{F}^\dagger\otimes\mathcal{F}^\#_t)_{t\geq 0},\mathbb{P}\otimes\mathbb{P}^*\otimes\mathbb{P}^\dagger\otimes\mathbb{P}^\#\big),$$

the four families $(\xi,X,\tilde X)$, $(\xi^*,X^*,\tilde X^*)$, $\xi^\dagger$ and W are independent. We consider the pathwise unique solutions U and $\tilde U$ to

$$\begin{aligned}U_t=&G(\xi,\xi^*,\xi^\dagger)+\int_0^t\alpha(X_s,X^*_s,U_s)\mathrm{d}s+\int_0^t\beta(X_s,X^*_s,U_s)\mathrm{d}W_s,\\ \tilde U_t=&G(\xi,\xi^*,\xi^\dagger)+\int_0^t\alpha(\tilde X_s,\tilde X^*_s,\tilde U_s)\mathrm{d}s+\int_0^t\beta(\tilde X_s,\tilde X^*_s,\tilde U_s)\mathrm{d}W_s.\end{aligned}$$

By Lemma 5, $U=\Gamma(\xi,X,\xi^*,X^*,\xi^\dagger,W)$ and $\tilde U=\Gamma(\xi,\tilde X,\xi^*,\tilde X^*,\xi^\dagger,W)$. Thus, as in (i),

$$\Lambda_{\xi,X}(t,\xi,X)=\mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\big[\phi(\Gamma_t(\xi,X,\xi^*,X^*,\xi^\dagger,W))\gamma(X_t,X^*_t)\big]\Big]\Big]=\mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\big[\phi(U_t)\gamma(X_t,X^*_t)\big]\Big]\Big].$$

Similarly, $\Lambda_{\xi,\tilde X}(t,\xi,\tilde X)=\mathbb{E}^*[\mathbb{E}^\dagger[\mathbb{E}^\#[\phi(\tilde U_t)\gamma(\tilde X_t,\tilde X^*_t)]]]$, so that

$$\begin{aligned}\mathbb{E}\Big[|\Lambda_{\xi,X}(t,\xi,X)-\Lambda_{\xi,\tilde X}(t,\xi,\tilde X)|^2\Big]=&\mathbb{E}\Big[\Big|\mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\phi(U_t)\gamma(X_t,X^*_t)-\phi(\tilde U_t)\gamma(\tilde X_t,\tilde X^*_t)\Big]\Big]\Big]\Big|^2\Big]\\ \leq&\mathbb{E}\Big[\mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\Big|\phi(U_t)\gamma(X_t,X^*_t)-\phi(\tilde U_t)\gamma(\tilde X_t,\tilde X^*_t)\Big|^2\Big]\Big]\Big]\Big]\\ =&\bar{\mathbb{E}}\Big[\Big|\phi(U_t)\gamma(X_t,X^*_t)-\phi(\tilde U_t)\gamma(\tilde X_t,\tilde X^*_t)\Big|^2\Big].\end{aligned}\tag{10}$$

Case 1: ϕ bounded. Using that ϕ,γ are bounded and Lipschitz continuous, (10) implies that

$$\mathbb{E}\Big[|\Lambda_{\xi,X}(t,\xi,X)-\Lambda_{\xi,\tilde X}(t,\xi,\tilde X)|^2\Big]\leq C\bar{\mathbb{E}}[|U_t-\tilde U_t|^2]+C\bar{\mathbb{E}}[|X_t-\tilde X_t|^2+|X^*_t-\tilde X^*_t|^2].\tag{11}$$

By Proposition 11-(ii),

$$\bar{\mathbb{E}}[|U_t-\tilde U_t|^2]\leq C_T\bar{\mathbb{E}}\Big[\sup_{s\in[0,t]}\big(|X_s-\tilde X_s|^2+|X^*_s-\tilde X^*_s|^2\big)\Big].\tag{12}$$

Inserting (12) into (11), we find

$$\mathbb{E}\Big[|\Lambda_{\xi,X}(t,\xi,X)-\Lambda_{\xi,\tilde X}(t,\xi,\tilde X)|^2\Big]\leq C_T\bar{\mathbb{E}}\Big[\sup_{s\in[0,t]}\big(|X_s-\tilde X_s|^2+|X^*_s-\tilde X^*_s|^2\big)\Big]\leq C_T\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s-\tilde X_s|^2\Big],$$

as desired, because $(X^*,\tilde X^*)$ has the same law as $(X,\tilde X)$.

Case 2: (6) holds. Since ϕ has linear growth, since ϕ,γ are Lipschitz continuous and since γ is bounded, (10) implies that

$$\begin{aligned}\mathbb{E}\Big[|\Lambda_{\xi,X}(t,\xi,X)-\Lambda_{\xi,\tilde X}(t,\xi,\tilde X)|^2\Big]\leq&C\bar{\mathbb{E}}[|U_t-\tilde U_t|^2]\\ &+C\bar{\mathbb{E}}\Big[(1+|U_t|^2)\big(|X_t-\tilde X_t|^2+|X^*_t-\tilde X^*_t|^2\big)\Big].\end{aligned}\tag{13}$$

Moreover, (12) still holds, while Proposition 11-(i) and (6) give, uniformly in $(\omega,\omega^*)\in\Omega\times\Omega^*$,

$$\mathbb{E}^\dagger\big[\mathbb{E}^\#[1+|U_t|^2]\big]\leq C_T.$$

Consequently,

$$
\begin{aligned}
&\bar{\mathbb{E}}\Big[(1+|U_t|^2)\big(|X_t-\tilde{X}_t|^2+|X_t^*-\tilde{X}_t^*|^2\big)\Big]\\
&=\mathbb{E}\Big[\mathbb{E}^*\Big[\big(|X_t-\tilde{X}_t|^2+|X_t^*-\tilde{X}_t^*|^2\big)\mathbb{E}^\dagger\big[\mathbb{E}^\#[1+|U_t|^2]\big]\Big]\Big]\\
&\le C_T\mathbb{E}\Big[\mathbb{E}^*\big[|X_t-\tilde{X}_t|^2+|X_t^*-\tilde{X}_t^*|^2\big]\Big]\\
&\le C_T\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s-\tilde{X}_s|^2\Big]. \qquad (14)
\end{aligned}
$$

The conclusion follows by inserting (12) and (14) into (13). ☐

2.2. **Proof of Theorem 8.**

Proof of Theorem 8. We first prove (i). We start with pathwise uniqueness and consider two solutions (ξ,X) and $(\xi,\tilde{X})$ to the nonlinear SDE, defined on the same probability space and driven by the same Brownian motion. Since b and σ are bounded and Lipschitz continuous, we obtain, for all $t\in[0,T]$,

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s-\tilde{X}_s|^2\Big] \le& C_T\int_0^t\mathbb{E}[|X_s-\tilde{X}_s|^2]\mathrm{d}s+C_T\int_0^t\mathbb{E}[|\Lambda_{\xi,X}(s,\xi,X)-\Lambda_{\xi,\tilde{X}}(s,\xi,\tilde{X})|^2]\mathrm{d}s\\
\le& C_T\int_0^t\mathbb{E}\Big[\sup_{u\in[0,s]}|X_u-\tilde{X}_u|^2\Big]\mathrm{d}u
\end{aligned}
$$

by Proposition 12-(ii). One completes the uniqueness proof using the Gronwall lemma.

We next check existence by a Picard iteration. We fix $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$ on which are defined ξ and B. We define $X^0=(X_t^0)_{t\ge0}$ by $X_t^0=F(\xi)$. Then X^0 satisfies (3) with $A_T=0$. Once $X^n=(X_t^n)_{t\ge0}$ (continuous, adapted and satisfying (3)) is built, we set

$$
X_t^{n+1}=F(\xi)+\int_0^t b(X_s^n)\mathrm{d}s+\int_0^t\sigma(X_s^n)\mathrm{d}B_s+\int_0^t\Lambda_{\xi,X^n}(s,\xi,X^n)\mathrm{d}s. \qquad (15)
$$

First, $X^{n+1}=(X_t^{n+1})_{t\ge0}$ is continuous and satisfies (3) for some constant A_T which does not depend on n, because b,σ and Λ_{ξ,X^n} are bounded, see Proposition 12-(i). Indeed, when computing, for $0\le s\le t\le s+1\le T+1$, the quantity $\mathbb{E}[|X_t^{n+1}-X_s^{n+1}|^2]$, the two drift terms are bounded in L^2 by $C_T(t-s)^2\le C_T(t-s)$, while Itô's isometry gives

$$
\mathbb{E}\Big[\Big|\int_s^t\sigma(X_r^n)\mathrm{d}B_r\Big|^2\Big]\le C_T(t-s).
$$

The same computation as in the uniqueness proof shows that for all $n\ge1$, all $t\in[0,T]$,

$$
\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s^{n+1}-X_s^n|^2\Big]\le C_T\int_0^t\mathbb{E}\Big[\sup_{u\in[0,s]}|X_u^n-X_u^{n-1}|^2\Big]\mathrm{d}u.
$$

The usual Picard argument implies that there is a continuous adapted process $X=(X_t)_{t\ge0}$ such that

$$
\lim_n\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t^n-X_t|^2\Big]=0. \qquad (16)
$$

Since X^n satisfies (3) for some constant A_T which does not depend on n, the limit X satisfies (3). Taking the limit $n\to\infty$ in (15), we find that $(X_t)_{t\ge0}$ solves the nonlinear SDE, because

$$
X_t^{n+1}\to X_t,\quad \int_0^t b(X_s^n)\mathrm{d}s\to\int_0^t b(X_s)\mathrm{d}s,\quad \int_0^t\sigma(X_s^n)\mathrm{d}B_s\to\int_0^t\sigma(X_s)\mathrm{d}B_s,
$$

$$
\text{and}\quad \int_0^t\Lambda_{\xi,X^n}(s,\xi,X^n)\mathrm{d}s\to\int_0^t\Lambda_{\xi,X}(s,\xi,X)\mathrm{d}s
$$

in $L^2(\Omega)$. The first three convergences follow from (16) and the Lipschitz continuity of b and σ, while the last one follows from Proposition 12-(ii) and (16).

We now prove (ii). The law of the solution (ξ, X) to the nonlinear SDE built above by Picard iteration does not depend on the choice of the probability space, nor on the choices of ξ and B. Indeed, one can check by induction that the law of (ξ, X^n, B) does not depend on these choices for all $n \geq 1$, because Λ_{ξ, X^n} is shorthand for $\Lambda_{\mathrm{Law}(\xi, X^n)}$, which depends only on the law of (ξ, X^n); see Lemma 6. The limit (ξ, X, B) inherits this property. In particular, the law of (ξ, X) does not depend on the probability space nor on the choices of ξ and B.

Consider now, on some space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, a solution $X = (X_t)_{t\geq 0}$ to the nonlinear SDE corresponding to some $\mathcal{F}_0$-measurable $\xi \sim \pi$ and to some m-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $B = (B_t)_{t\geq 0}$. On this space, consider the solution $(\xi, \hat{X})$ built by Picard iteration with ξ and B. By the pathwise uniqueness proved in (i), $(\xi, X) = (\xi, \hat{X})$ a.s. The conclusion follows. □

2.3. **Proof of Theorem 9.**

Proof of Theorem 9. We first prove (i). Recall that $(X_t^{i,N})_{t\geq 0, i\in I_N}$ and $(U_t^{ij,N})_{t\geq 0, i,j\in I_N, i\neq j}$ were defined in (1)-(2). For all $i \geq 1$, we denote by $(\bar{X}_t^i)_{t\geq 0}$ the solution to the nonlinear SDE corresponding to ξ_i and B^i, i.e.

$$\bar{X}_t^i = F(\xi_i) + \int_0^t b(\bar{X}_s^i)\mathrm{d}s + \int_0^t \sigma(\bar{X}_s^i)\mathrm{d}B_s^i + \int_0^t \Lambda_{\xi_i, \bar{X}^i}(s, \xi_i, \bar{X}^i)\mathrm{d}s, \tag{17}$$

and, for $i, j \geq 1$ with $i \neq j$, we consider $(\bar{U}_t^{ij})_{t\geq 0}$ solving (9). As a strong solution, $\bar{X}^i$ is $\sigma(\xi_i, B^i)$-measurable. Using moreover Theorem 8-(ii) and Assumption 1, we conclude that

(18) the family $((\xi_i, \bar{X}^i), i \geq 1)$ is i.i.d. and independent of $((\xi_{ij}^\dagger, W^{ij}), i, j \geq 1, i \neq j)$.

Since α and β are bounded and Lipschitz continuous, we have, for all $t \in [0, T]$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,t]} |U_s^{12,N} - \bar{U}_s^{12}|^2\Big] \leq& C_T \int_0^t \mathbb{E}[|X_s^{1,N} - \bar{X}_s^1|^2 + |X_s^{2,N} - \bar{X}_s^2|^2 + |U_s^{12,N} - \bar{U}_s^{12}|^2]\mathrm{d}s \\ \leq& C_T \int_0^t \mathbb{E}[|X_s^{1,N} - \bar{X}_s^1|^2 + |U_s^{12,N} - \bar{U}_s^{12}|^2]\mathrm{d}s\end{aligned}$$

by exchangeability. Thanks to the Gronwall lemma, for all $t \in [0, T]$,

$$\mathbb{E}\Big[\sup_{s\in[0,t]} |U_s^{12,N} - \bar{U}_s^{12}|^2\Big] \leq C_T \int_0^t \mathbb{E}[|X_s^{1,N} - \bar{X}_s^1|^2]\mathrm{d}s \leq C_T \mathbb{E}\Big[\sup_{s\in[0,t]} |X_s^{1,N} - \bar{X}_s^1|^2\Big]. \tag{19}$$

Since b and σ are Lipschitz continuous, for all $t \in [0, T]$,

$$\mathbb{E}\Big[\sup_{s\in[0,t]} |X_s^{1,N} - \bar{X}_s^1|^2\Big] \leq C_T \int_0^t \mathbb{E}[|X_s^{1,N} - \bar{X}_s^1|^2]\mathrm{d}s + C_T \int_0^t I_s \mathrm{d}s, \tag{20}$$

where, setting $\mu = \mathrm{Law}(\xi_1, \bar{X}^1)$,

$$I_t = \mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^N \phi(U_t^{1j,N})\gamma(X_t^{1,N}, X_t^{j,N}) - \Lambda_\mu(t, \xi_1, \bar{X}^1)\Big|^2\Big].$$

We write $I_t \le 2J_t + 2K_t$, where J_t is the stability error and K_t is the empirical fluctuation error:

$$J_t = \mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^{N}\Big(\phi(U_t^{1j,N})\gamma(X_t^{1,N},X_t^{j,N}) - \phi(\bar U_t^{1j})\gamma(\bar X_t^1,\bar X_t^j)\Big)\Big|^2\Big],$$
$$K_t = \mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^{N}\phi(\bar U_t^{1j})\gamma(\bar X_t^1,\bar X_t^j) - \Lambda_\mu(t,\xi_1,\bar X^1)\Big|^2\Big].$$

By the Cauchy–Schwarz inequality and since ϕ and γ are bounded and Lipschitz continuous,

$$\begin{aligned}
J_t \le& \frac{1}{N-1}\sum_{j=2}^{N}\mathbb{E}\Big[\Big|\phi(U_t^{1j,N})\gamma(X_t^{1,N},X_t^{j,N}) - \phi(\bar U_t^{1j})\gamma(\bar X_t^1,\bar X_t^j)\Big|^2\Big]\\
\le& \frac{C}{N-1}\sum_{j=2}^{N}\mathbb{E}\Big[|U_t^{1j,N}-\bar U_t^{1j}|^2 + |X_t^{1,N}-\bar X_t^1|^2 + |X_t^{j,N}-\bar X_t^j|^2\Big]\\
\le& \frac{C_T}{N-1}\sum_{j=2}^{N}\mathbb{E}\Big[\sup_{s\in[0,t]}(|X_s^{1,N}-\bar X_s^1|^2 + |X_s^{j,N}-\bar X_s^j|^2)\Big]\\
\le& C_T\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s^{1,N}-\bar X_s^1|^2\Big]
\end{aligned}$$

by (19) and exchangeability. Since $\bar U_t^{1j} = \Gamma_t(\xi_1,\bar X^1,\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j})$ by Lemma 5,

$$K_t = \mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^{N}\phi(\Gamma_t(\xi_1,\bar X^1,\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j}))\gamma(\bar X_t^1,\bar X_t^j) - \Lambda_\mu(t,\xi_1,\bar X^1)\Big|^2\Big].$$

By independence of $(\xi_1,\bar X^1)$ and $(\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j})_{j\ge2}$, see (18), The family

$$\big(\phi(\Gamma_t(\xi_1,\bar X^1,\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j}))\gamma(\bar X_t^1,\bar X_t^j)\big)_{j\ge2}$$

is i.i.d. conditionally on $(\xi_1,\bar X^1)$. Moreover, by definition of Λ_μ,

$$\mathbb{E}_1\Big[\phi(\Gamma_t(\xi_1,\bar X^1,\xi_2,\bar X^2,\xi_{12}^\dagger,W^{12}))\gamma(\bar X_t^1,\bar X_t^2)\Big] = \Lambda_\mu(t,\xi_1,\bar X^1),$$

where $\mathbb{E}_1 = \mathbb{E}[\cdot|(\xi_1,\bar X^1)]$. Hence, abusively writing $\mathrm{Var}_1 Z = \mathbb{E}_1[|Z-\mathbb{E}_1[Z]|^2]$ for Z valued in $\mathbb{R}^d$,

$$K_t = \mathbb{E}\Big[\frac{1}{N-1}\mathrm{Var}_1\Big(\phi(\Gamma_t(\xi_1,\bar X^1,\xi_2,\bar X^2,\xi_{12}^\dagger,W^{12}))\gamma(\bar X_t^1,\bar X_t^2)\Big)\Big] \le \frac{||\phi||_\infty^2||\gamma||_\infty^2}{N-1} \le \frac{C}{N}.$$

We used that $N\ge2$ in the last inequality.

All in all, we have proved that for all $t\in[0,T]$,

$$I_t \le 2J_t + 2K_t \le C_T\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s^{1,N}-\bar X_s^1|^2\Big] + \frac{C}{N}.$$

Recalling (20), we get, for all $t\in[0,T]$,

$$\mathbb{E}\Big[\sup_{s\in[0,t]}|X_s^{1,N}-\bar X_s^1|^2\Big] \le C_T\int_0^t \mathbb{E}\Big[\sup_{u\in[0,s]}|X_u^{1,N}-\bar X_u^1|^2\Big]\mathrm{d}s + \frac{C_T}{N}.$$

Using finally the Gronwall lemma, we find as desired that

$$\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t^{1,N}-\bar X_t^1|^2\Big] \le \frac{C_T}{N}, \quad \text{whence also} \quad \mathbb{E}\Big[\sup_{t\in[0,T]}|U_t^{12,N}-\bar U_t^{12}|^2\Big] \le \frac{C_T}{N}$$

by (19).

We now prove (ii). We adopt the same notation as above and observe that we still have (18). We fix $T>0$ and divide the proof into several steps.

Step 1 : Gaussian moment for the edge. There are $\theta_T>0$ and $C_T>0$ such that

$$\mathbb{E}\Big[\exp\Big(2\theta_T\sup_{t\in[0,T]}(\bar U_t^{12})^2\Big)\Big]\le C_T.$$

Recall (9): setting $G_{12}=G(\xi_1,\xi_2,\xi_{12}^\dagger)$ and $R_t=\int_0^t\beta(\bar X_s^1,\bar X_s^2,\bar U_s^{12})\mathrm{d}W_s^{12}$,

$$\sup_{t\in[0,T]}|\bar U_t^{12}|\le|G_{12}|+T||\alpha||_\infty+R_T^*,\quad\text{where}\quad R_T^*=\sup_{t\in[0,T]}|R_t|.$$

By (8), G_{12} has a Gaussian moment. By the Dubins-Schwarz theorem, see e.g. Revuz-Yor [29, Theorem 1.6 p 181], we can find a Brownian motion M such that a.s., for all $t\ge0$, $R_t=M_{A_t}$, with $A_t=\int_0^t\beta^2(\bar X_s^1,\bar X_s^2,\bar U_s^{12})\mathrm{d}s$. Thus $R_T^*\le\sup_{t\in[0,||\beta||_\infty^2T]}|M_t|$. Hence R_T^* has a Gaussian moment. Choosing $\theta_T>0$ small enough and using $(x+y+z)^2\le3(x^2+y^2+z^2)$ together with the Cauchy–Schwarz inequality, we get the conclusion.

Step 2: Empirical fluctuation. Set $\mu=\mathrm{Law}(\xi_1,\bar X^1)$. There is $C_T>0$ such that for all $t\in[0,T]$,

$$\Delta_t:=\mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^N\phi(\bar U_t^{1j})\gamma(\bar X_t^1,\bar X_t^j)-\Lambda_\mu(t,\xi_1,\bar X^1)\Big|\Big]\le\frac{C_T}{N^{\frac12}}.$$

As in the proof of (i) (study of K_t) we have $\Delta_t=\mathbb{E}[E_t]$, where, setting $\mathbb{E}_1=\mathbb{E}[\cdot|(\xi_1,\bar X^1)]$,

$$E_t=\mathbb{E}_1\Big[\Big|\frac{1}{N-1}\sum_{j=2}^N\phi(\Gamma_t(\xi_1,\bar X^1,\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j}))\gamma(\bar X_t^1,\bar X_t^j)-\Lambda_\mu(t,\xi_1,\bar X^1)\Big|\Big].$$

We have

$$\begin{aligned}E_t\le&\mathbb{E}_1\Big[\Big|\frac{1}{N-1}\sum_{j=2}^N\phi(\Gamma_t(\xi_1,\bar X^1,\xi_j,\bar X^j,\xi_{1j}^\dagger,W^{1j}))\gamma(\bar X_t^1,\bar X_t^j)-\Lambda_\mu(t,\xi_1,\bar X^1)\Big|^2\Big]^{\frac12}\\\le&\Big(\frac{1}{N-1}\mathrm{Var}_1\Big(\phi(\Gamma_t(\xi_1,\bar X^1,\xi_2,\bar X^2,\xi_{12}^\dagger,W^{12}))\gamma(\bar X_t^1,\bar X_t^2)\Big)\Big)^{\frac12}\end{aligned}$$

as in the proof of (i). Thus, recalling that $\Gamma_t(\xi_1,\bar X^1,\xi_2,\bar X^2,\xi_{12}^\dagger,W^{12})=\bar U_t^{12}$, for all $N\ge2$,

$$\Delta_t\le\sqrt{\frac{2}{N}}\mathbb{E}[(\phi(\bar U_t^{12})\gamma(\bar X_t^1,\bar X_t^2))^2]\le\frac{C}{N^{\frac12}}\mathbb{E}[1+(\bar U_t^{12})^2]\le\frac{C_T}{N^{\frac12}}.$$

We used that ϕ has at most linear growth, that γ is bounded, and that $\sup_{[0,T]}\mathbb{E}[(\bar U_t^{12})^2]<\infty$.

Step 3 : Notation. For $i,j\in I_N$ with $i\ne j$ and $0\le s\le t$, set

$$Y_{s,t}^{i,N}=\sup_{u\in[s,t]}|X_u^{i,N}-\bar X_u^i|\quad\text{and}\quad Z_{s,t}^{ij,N}=\sup_{u\in[s,t]}|U_u^{ij,N}-\bar U_u^{ij}|.$$

Step 4. Edge stability. In this step we establish a short-time estimate for the edge error $Z^{12,N}$. Fix $0\le S\le S'\le T$ and $t\in[S,S']$. Recalling (2) and (9), we see that $Z_{S,t}^{12,N}\le Z_{S,S}^{12,N}+I_{S,t}^1+I_{S,t}^2$, where

$$\begin{aligned}I_{S,t}^1&=\int_S^t|\alpha(X_s^{1,N},X_s^{2,N},U_s^{12,N})-\alpha(\bar X_s^1,\bar X_s^2,\bar U_s^{12})|\mathrm{d}s,\\I_{S,t}^2&=\sup_{r\in[S,t]}\Big|\int_S^r(\beta(X_s^{1,N},X_s^{2,N},U_s^{12,N})-\beta(\bar X_s^1,\bar X_s^2,\bar U_s^{12}))\mathrm{d}W_s^{12}\Big|.\end{aligned}$$

Using that α is Lipschitz continuous, we find

$$\mathbb{E}[I_{S,t}^1]\le C(t-S)\mathbb{E}[Y_{S,t}^{1,N}+Y_{S,t}^{2,N}+Z_{S,t}^{12,N}]\le C(S'-S)\mathbb{E}[Y_{S,t}^{1,N}+Z_{S,t}^{12,N}]$$

by exchangeability. Using the Burkholder–Davis–Gundy inequality and that β is Lipschitz continuous,

$$\mathbb{E}[I^2_{S,t}] \le C\mathbb{E}\Big[\Big(\int_S^t (\beta(X^{1,N}_s, X^{2,N}_s, U^{12,N}_s) - \beta(\bar{X}^1_s, \bar{X}^2_s, \bar{U}^{12}_s))^2 \mathrm{d}s\Big)^{\frac{1}{2}}\Big] \le C(S'-S)^{\frac{1}{2}}\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}]$$

by exchangeability again. Combining the bounds on $I^1_{S,t}$ and $I^2_{S,t}$, we get, for all $0 \le S \le t \le S' \le T$,

$$\mathbb{E}[Z^{12,N}_{S,t}] \le \mathbb{E}[Z^{12,N}_{S,S}] + \kappa_1[S'-S+(S'-S)^{\frac{1}{2}}]\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}]$$

for some $\kappa_1 > 0$ depending only on α and β.

Step 5: Particle stability. Here we establish a short-time estimate for $Y^{1,N}$. Fix $0 \le S \le S' \le T$ and $t \in [S, S']$. Recalling (1) and (17), we have $Y^{1,N}_{S,t} \le Y^{1,N}_{S,S} + J^1_{S,t} + J^2_{S,t} + J^3_{S,t}$, where

$$J^1_{S,t} = \int_S^t |b(X^{1,N}_s) - b(\bar{X}^1_s)|\mathrm{d}s, \quad J^2_{S,t} = \sup_{r\in[S,t]}\Big|\int_S^r (\sigma(X^{1,N}_s) - \sigma(\bar{X}^1_s))\mathrm{d}B^1_s\Big|,$$

$$J^3_{S,t} = \int_S^t \Big|\frac{1}{N-1}\sum_{j=2}^N \phi(U^{1j,N}_s)\gamma(X^{1,N}_s, X^{j,N}_s) - \Lambda_\mu(s,\xi_1,\bar{X}^1)\Big|\mathrm{d}s.$$

The terms $J^1_{S,t}$ and $J^2_{S,t}$ are the drift and martingale contributions coming from the single-particle coefficients b and σ, while $J^3_{S,t}$ is the interaction contribution. Exactly as in Step 4, using the Lipschitz continuity of b and σ and the Burkholder–Davis–Gundy inequality, we find that

$$\mathbb{E}[J^1_{S,t} + J^2_{S,t}] \le C[S'-S+(S'-S)^{\frac{1}{2}}]\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}].$$

Recalling Step 2, we write

$$\begin{aligned}\mathbb{E}[J^3_{S,t}] &\le \int_S^t \mathbb{E}\Big[\Big|\frac{1}{N-1}\sum_{j=2}^N \Big(\phi(U^{1j,N}_s)\gamma(X^{1,N}_s, X^{j,N}_s) - \phi(\bar{U}^{1j}_s)\gamma(\bar{X}^1_s, \bar{X}^j_s)\Big)\Big|\Big]\mathrm{d}s + \int_S^t \Delta_s \mathrm{d}s \\ &\le \int_S^t \mathbb{E}[|\phi(U^{12,N}_s)\gamma(X^{1,N}_s, X^{2,N}_s) - \phi(\bar{U}^{12}_s)\gamma(\bar{X}^1_s, \bar{X}^2_s)|]\mathrm{d}s + C_T N^{-\frac{1}{2}}\end{aligned}$$

by exchangeability and Step 2. We now set $H^N_s = (|X^{1,N}_s - \bar{X}^1_s| \wedge 1) + (|X^{2,N}_s - \bar{X}^2_s| \wedge 1)$. Since γ and ϕ are Lipschitz continuous, with γ bounded, for $t \in [S, S']$,

$$\begin{aligned}\mathbb{E}[J^3_{S,t}] \le& C\int_S^t \mathbb{E}[(1+|\bar{U}^{12}_s|)H^N_s + |U^{12,N}_s - \bar{U}^{12}_s|]\mathrm{d}s + C_T N^{-\frac{1}{2}} \\ \le& C\int_S^t \mathbb{E}[|\bar{U}^{12}_s|H^N_s]\mathrm{d}s + C(S'-S)\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] + C_T N^{-\frac{1}{2}}.\end{aligned}$$

Combining these estimates (and using exchangeability), we get, for all $0 \le S \le t \le S' \le T$,

$$\mathbb{E}[Y^{1,N}_{S,t}] \le \mathbb{E}[Y^{1,N}_{S,S}] + C_T N^{-\frac{1}{2}} + \kappa_2[S'-S+(S'-S)^{\frac{1}{2}}]\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] + \kappa_2\int_S^t \mathbb{E}[|\bar{U}^{12}_s|H^N_s]\mathrm{d}s$$

for some $\kappa_2 > 0$ depending only on b, σ, γ and ϕ.

Step 6: Truncation and recursion. Gathering Steps 4 and 5, we find $\kappa_3 > 0$ such that for all $0 \le S \le S' \le T$, all $t \in [S, S']$,

$$\begin{aligned}\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] \le& \mathbb{E}[Y^{1,N}_{S,S} + Z^{12,N}_{S,S}] + C_T N^{-\frac{1}{2}} + \kappa_3[S'-S+(S'-S)^{\frac{1}{2}}]\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] \\ &+ \kappa_3\int_S^t \mathbb{E}[|\bar{U}^{12}_s|H^N_s]\mathrm{d}s.\end{aligned}$$

Let $\delta > 0$ such that $\kappa_3(\delta + \delta^{\frac{1}{2}}) = \frac{1}{2}$ and $\kappa_4 = 2\kappa_3$. Then for all $0 \le S \le S' \le T \wedge (S+\delta)$, all $t \in [S, S']$,

$$\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] \le 2\mathbb{E}[Y^{1,N}_{S,S} + Z^{12,N}_{S,S}] + C_T N^{-\frac{1}{2}} + \kappa_4 \int_S^t \mathbb{E}[|\bar{U}^{12}_s| H^N_s]\mathrm{d}s.$$

The only remaining difficulty is the unbounded factor $\bar{U}^{12}$ in the last integral. We handle it by truncating at a suitable level A. For any $A > 0$ and $s \ge S$, using exchangeability and Step 1,

$$\begin{aligned}\mathbb{E}[|\bar{U}^{12}_s| H^N_s] &\le A\mathbb{E}[|X^{1,N}_s - \bar{X}^1_s| + |X^{2,N}_s - \bar{X}^2_s|] + 2\mathbb{E}[|\bar{U}^{12}_s| \mathbf{1}_{\{|\bar{U}^{12}_s| > A\}}]\\ &\le 2A\mathbb{E}[Y^{1,N}_{S,s}] + C_T e^{-\theta_T A^2}.\end{aligned}$$

In the last line, we used that $u\mathbf{1}_{\{u>A\}} \le C_T e^{-\theta_T A^2} e^{2\theta_T u^2}$ and the Gaussian moment from Step 1. Plugging this into the preceding inequality, we get that for all $0 \le S \le S' \le T \wedge (S+\delta)$, all $t \in [S, S']$ and all $A > 0$,

$$\mathbb{E}[Y^{1,N}_{S,t} + Z^{12,N}_{S,t}] \le 2\mathbb{E}[Y^{1,N}_{S,S} + Z^{12,N}_{S,S}] + C_T(N^{-\frac{1}{2}} + e^{-\theta_T A^2}) + 2A\kappa_4 \int_S^t \mathbb{E}[Y^{1,N}_{S,s} + Z^{12,N}_{S,s}]\mathrm{d}s.$$

Using the Gronwall lemma, we find that for all $0 \le S \le S' \le T \wedge (S+\delta)$, for all $A > 0$,

$$\mathbb{E}[Y^{1,N}_{S,S'} + Z^{12,N}_{S,S'}] \le \Big(2\mathbb{E}[Y^{1,N}_{S,S} + Z^{12,N}_{S,S}] + C_T(N^{-\frac{1}{2}} + e^{-\theta_T A^2})\Big) e^{2\kappa_4 A\delta}.$$

Choosing $A = (\frac{\log N}{2\theta_T})^{\frac{1}{2}}$, this gives, for all $0 \le S \le S' \le T \wedge (S+\delta)$,

$$\begin{aligned}&\mathbb{E}[Y^{1,N}_{S,S'} + Z^{12,N}_{S,S'}] \le \Big(2\mathbb{E}[Y^{1,N}_{S,S} + Z^{12,N}_{S,S}] + C_T N^{-\frac{1}{2}}\Big) K_{N,T},\\ &\text{where} \quad K_{N,T} = \exp\Big(2\kappa_4\delta\Big(\frac{\log N}{2\theta_T}\Big)^{\frac{1}{2}}\Big).\end{aligned} \tag{21}$$

Step 7: Conclusion. Set $k = \lceil T/\delta \rceil$ and $t_\ell = \ell\delta \wedge T$, for $\ell = 0, \dots, k$. Set $u_{N,0} = 0$ and, for $\ell = 0, \dots, k-1$,

$$u_{N,\ell+1} = \mathbb{E}[Y^{1,N}_{t_\ell, t_{\ell+1}} + Z^{12,N}_{t_\ell, t_{\ell+1}}].$$

By (21), for all $\ell = 0, \dots, k-1$,

$$u_{N,\ell+1} \le (2u_{N,\ell} + C_T N^{-\frac{1}{2}}) K_{N,T}.$$

Indeed, the error at the left endpoint of the ℓ-th block is bounded by $u_{N,\ell}$. We classically conclude that, modifying the value of C_T,

$$u_{N,\ell} \le C_T K^k_{N,T} N^{-\frac{1}{2}}$$

for all $\ell = 1, \dots, k$. All in all, modifying again the value of C_T,

$$\mathbb{E}\Big[\sup_{s\in[0,T]} (|X^{1,N}_s - \bar{X}^1_s| + |U^{12,N}_s - \bar{U}^{12}_s|)\Big] \le \sum_{\ell=1}^k u_{N,\ell} \le C_T K^k_{N,T} N^{-\frac{1}{2}}.$$

Since

$$K^k_{N,T} = \exp\Big(2\kappa_4 k\delta\Big(\frac{\log N}{2\theta_T}\Big)^{\frac{1}{2}}\Big),$$

the conclusion of (ii) follows.

We finally prove (iii). Under the standing alternative in the statement, either the assumption of Theorem 9-(i) or that of Theorem 9-(ii) holds. We let $P_N := \{(i,j) \in I_N, i \ne j\}$. As already seen, $\bar{X}^i$ is $\sigma(\xi_i, B^i)$-measurable and $\bar{U}^{ij}$ is $\sigma(\xi_i, B^i, \xi_j, B^j, \xi^\dagger_{ij}, W^{ij})$-measurable. Using Theorem 8 and Assumption 1, we see that all the random variables $\mathcal{Z}^{ij}_t = (\bar{X}^i_t, \bar{X}^j_t, \bar{U}^{ij}_t)$ are f_t-distributed and that for $C \subset P_N$, the family $(\mathcal{Z}_t^{ij}, (i,j) \in C)$ is i.i.d. as soon as for all $(i,j), (i',j') \in C$ distinct, the four numbers i, j, i', j' are distinct. We set

$$\mu_t^N = \frac{1}{N(N-1)} \sum_{(i,j)\in P_N} \delta_{(X_t^{i,N}, X_t^{j,N}, U_t^{ij,N})} \quad \text{and} \quad \bar{\mu}_t^N = \frac{1}{N(N-1)} \sum_{(i,j)\in P_N} \delta_{(\bar{X}_t^i, \bar{X}_t^j, \bar{U}_t^{ij})}.$$

We write $\mathcal{W}_1(\mu_t^N, f_t) \le \mathcal{W}_1(\mu_t^N, \bar{\mu}_t^N) + \mathcal{W}_1(\bar{\mu}_t^N, f_t)$. We classically have

$$\mathbb{E}[\mathcal{W}_1(\mu_t^N, \bar{\mu}_t^N)] \le \frac{1}{N(N-1)} \sum_{(i,j)\in P_N} \mathbb{E}[|X_t^{i,N} - \bar{X}_t^i| + |X_t^{j,N} - \bar{X}_t^j| + |U_t^{ij,N} - \bar{U}_t^{ij}|] \le \frac{C_T}{N^{\frac{1}{2d+1}}}$$

by Theorem 9-(i) (since $N^{-\frac{1}{2}} \le N^{-\frac{1}{2d+1}}$) or Theorem 9-(ii) (since $N^{-\frac{1}{2}} \exp(C_T\sqrt{\log N}) \le C_T N^{-\frac{1}{2d+1}}$), using Cauchy–Schwarz in the first case to pass from the L^2 estimate to L^1.

We next write $P_N = \bigsqcup_{k=0}^{2N-1} P_N^k$, where, for $k = 0, \dots, N-1$,

$$P_N^k = \{(i,j) \in I_N : i + j = k \mod N, i < j\} \text{ and } P_N^{N+k} = \{(i,j) \in I_N : i + j = k \mod N, i > j\}.$$

Easy considerations show that for all $k = 0, \dots, 2N-1$, all $(i,j), (i',j') \in P_N^k$ with $(i,j) \neq (i',j')$, the four numbers i, i', j, j' are distinct, so that the family $(\mathcal{Z}_t^{ij}, (i,j) \in P_N^k)$ is i.i.d. Moreover,

$$\#(P_N^k) = \begin{cases} \frac{N-1}{2} & \text{if } N \text{ is odd,} \\ \frac{N}{2} & \text{if } N \text{ is even and } k \text{ is odd,} \\ \frac{N}{2} - 1 & \text{if } N \text{ is even and } k \text{ is even.} \end{cases}$$

Assuming now that $N \ge 3$ is odd, we have

$$\bar{\mu}_t^N = \frac{1}{2N} \sum_{k=0}^{2N-1} \bar{\mu}_t^{N,k}, \quad \text{where} \quad \bar{\mu}_t^{N,k} = \frac{2}{N-1} \sum_{(i,j)\in P_N^k} \delta_{(\bar{X}_t^i, \bar{X}_t^j, \bar{U}_t^{ij})}.$$

Using the joint convexity of $\mathcal{W}_1$ in its measure arguments, we have

$$\mathcal{W}_1(\bar{\mu}_t^N, f_t) \le \frac{1}{2N} \sum_{k=0}^{2N-1} \mathcal{W}_1(\bar{\mu}_t^{N,k}, f_t).$$

For each $k = 0, \dots, 2N-1$, $\bar{\mu}_t^{N,k}$ is the empirical measure of $\frac{N-1}{2}$ i.i.d. random variables with law $f_t \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R})$. Using the defining equations, Assumption 2, the standing alternative in the theorem, and the boundedness of the coefficients, one easily checks that for all $T > 0$,

$$\sup_{t\in[0,T]} \mathbb{E}[|\bar{X}_t^1|^2 + |\bar{X}_t^2|^2 + |\bar{U}_t^{12}|^2] < \infty.$$

By [12, Theorem 1] with (p,d,q) replaced by $(1, 2d+1, 2)$, for all $t \in [0,T]$, all $k = 0, \dots, 2N-1$,

$$\mathbb{E}[\mathcal{W}_1(\bar{\mu}_t^{N,k}, f_t)] \le C_T N^{-\frac{1}{2d+1}}.$$

All this shows that if $N \ge 3$ is odd,

$$\mathbb{E}[\mathcal{W}_1(\bar{\mu}_t^N, f_t)] \le C_T N^{-\frac{1}{2d+1}}, \quad \text{whence} \quad \mathbb{E}[\mathcal{W}_1(\mu_t^N, f_t)] \le C_T N^{-\frac{1}{2d+1}}.$$

The case where $N \ge 4$ is even is treated similarly, with light complications. Taking the supremum over $t \in [0,T]$ gives the claim. □

3. Measurability

We first prove Lemma 5, using arguments found in Karandikar [26].

Proof of Lemma 5. We set $\mathcal{X} = E \times \mathcal{C}_d \times E \times \mathcal{C}_d \times E^\dagger \times \mathcal{C}_1$.

Step 1: Euler scheme. Fix $n \geq 1$. For $(a, x, b, y, c, w) \in \mathcal{X}$, we denote by $(\Gamma^n_t(a, x, b, y, c, w))_{t\geq 0}$ the Euler scheme with step 2^{-n} for the equation $u_t = G(a, b, c) + \int_0^t \alpha(x_s, y_s, u_s)\mathrm{d}s + \int_0^t \beta(x_s, y_s, u_s)\mathrm{d}w_s$: we set $q_0 = G(a, b, c)$ and, for $k \geq 0$,

$$q_{(k+1)2^{-n}} = q_{k2^{-n}} + 2^{-n}\alpha(x_{k2^{-n}}, y_{k2^{-n}}, q_{k2^{-n}}) + (w_{(k+1)2^{-n}} - w_{k2^{-n}})\beta(x_{k2^{-n}}, y_{k2^{-n}}, q_{k2^{-n}}).$$

Finally, for all $t \geq 0$, choose $k = \lfloor 2^n t\rfloor$ and set

$$\Gamma^n_t(a, x, b, y, c, w) = q_{k2^{-n}} + \Big(t - k2^{-n}\Big)\alpha(x_{k2^{-n}}, y_{k2^{-n}}, q_{k2^{-n}}) + (w_t - w_{k2^{-n}})\beta(x_{k2^{-n}}, y_{k2^{-n}}, q_{k2^{-n}}).$$

The map $\Gamma^n : \mathcal{X} \to \mathcal{C}_1$ is measurable, because it is a continuous function of $(G(a, b, c), x, y, w)$. By construction, it holds that for all $(a, x, b, y, c, w) \in \mathcal{X}$, all $t \geq 0$,

$$\Gamma^n_t(a, x, b, y, c, w) = \Gamma^n_t(a, (x_{s\wedge t})_{s\geq 0}, b, y, c, w). \tag{22}$$

Step 2: Convergence to the edge SDE. Here we prove that for any space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, any independent $\mathcal{F}_0$-measurable triple $(\xi, \xi', \xi^\dagger)$ with $\xi, \xi' \sim \pi$ and $\xi^\dagger \sim \pi^\dagger$, any pair of continuous $(\mathcal{F}_t)_{t\geq 0}$-adapted processes $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ valued in $\mathbb{R}^d$ both satisfying (3), any 1-dimensional $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $W = (W_t)_{t\geq 0}$, the pathwise unique solution $U = (U_t)_{t\geq 0}$ to (4) satisfies

$$\text{for all } T > 0, \quad \lim_n \sup_{t\in[0,T]} |U_t - \Gamma^n_t(\xi, X, \xi', Y, \xi^\dagger, W)|^2 = 0 \quad \text{a.s.} \tag{23}$$

Since α and β are bounded, there is a constant $C > 0$ such that for all $0 \leq s < t < s + 1$,

$$\mathbb{E}[|U_t - U_s|^2] \leq C(t - s). \tag{24}$$

We next set $U^n_t = \Gamma^n_t(\xi, X, \xi', Y, \xi^\dagger, W)$, which classically satisfies, introducing $\rho_n(t) = 2^{-n}\lfloor 2^n t\rfloor$,

$$U^n_t = G(\xi, \xi', \xi^\dagger) + \int_0^t \alpha(X_{\rho_n(s)}, Y_{\rho_n(s)}, U^n_{\rho_n(s)})\mathrm{d}s + \int_0^t \beta(X_{\rho_n(s)}, Y_{\rho_n(s)}, U^n_{\rho_n(s)})\mathrm{d}W_s.$$

Since α and β are bounded and Lipschitz continuous, we have, for all $t \in [0, T]$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,t]} |U_s - U^n_s|^2\Big] \leq & C_T \int_0^t \mathbb{E}\Big[|X_s - X_{\rho_n(s)}|^2 + |Y_s - Y_{\rho_n(s)}|^2 + |U_s - U^n_{\rho_n(s)}|^2\Big]\mathrm{d}s \\ \leq & C_T\varepsilon_{T,n} + C_T \int_0^t \mathbb{E}\Big[\sup_{u\in[0,s]} |U_u - U^n_u|^2\Big]\mathrm{d}s,\end{aligned}$$

where

$$\varepsilon_{T,n} := \int_0^T \mathbb{E}\Big[|X_s - X_{\rho_n(s)}|^2 + |Y_s - Y_{\rho_n(s)}|^2 + |U_s - U_{\rho_n(s)}|^2\Big]\mathrm{d}s \leq C_T 2^{-n}$$

by (24), since X and Y satisfy (3) and since $0 \leq s - \rho_n(s) \leq 2^{-n}$ for all $s \geq 0$. By Gronwall's lemma,

$$\mathbb{E}\Big[\sup_{s\in[0,T]} |U_s - U^n_s|^2\Big] \leq C_T 2^{-n}.$$

Thus the series $\sum_{n\geq 1} \sup_{s\in[0,T]} |U_s - U^n_s|^2$ a.s. converges, whence $\lim_n \sup_{s\in[0,T]} |U_s - U^n_s|^2 = 0$ a.s. Taking the intersection of the corresponding probability-one events over integer $T \geq 1$, we also have $U^n \to U$ in $(\mathcal{C}_1, \delta)$ a.s. where $\delta(z, z') = \sum_{\ell\geq 1} 2^{-\ell}[1 \wedge \sup_{t\in[0,\ell]} |z(t) - z'(t)|]$ classically metrizes the topology of uniform convergence on compact time intervals.

Step 3: Measurability of the limit map. Since $(\mathcal{C}_1, \delta)$ is complete, the set

$$\mathcal{A} = \{(a, x, b, y, c, w) \in \mathcal{X} : \lim_n \Gamma^n(a, x, b, y, c, w) \text{ exists in } (\mathcal{C}_1, \delta)\}$$

is measurable. We then set

$$\Gamma(a,x,b,y,c,w) = \lim_n \mathbf{1}_{\{(a,x,b,y,c,w)\in\mathcal{A}\}}\Gamma^n(a,x,b,y,c,w)$$

for all $(a,x,b,y,c,w)\in\mathcal{X}$. This map $\Gamma:\mathcal{X}\to\mathcal{C}_1$ is measurable.

Step 4: Identification with the SDE. Consider a space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq0},\mathbb{P})$, a pair of independent π-distributed $\mathcal{F}_0$-measurable random variables ξ,ξ' valued in E, an $\mathcal{F}_0$-measurable $E^\dagger$-valued random variable $\xi^\dagger$ independent of (ξ,ξ') and with law $\pi^\dagger$, a pair of continuous $(\mathcal{F}_t)_{t\geq0}$-adapted processes $X=(X_t)_{t\geq0}$ and $Y=(Y_t)_{t\geq0}$ valued in $\mathbb{R}^d$ both satisfying (3), a 1-dimensional $(\mathcal{F}_t)_{t\geq0}$-Brownian motion $W=(W_t)_{t\geq0}$, and the pathwise unique solution $U=(U_t)_{t\geq0}$ to (4). Consider the event $A=\{(\xi,X,\xi',Y,\xi^\dagger,W)\in\mathcal{A}\}$, which has probability 1 by Step 2. On A, we have $U=\Gamma(\xi,X,\xi',Y,\xi^\dagger,W)$ (in the sense that for all $t\geq0$, $U_t=\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W)$) again by Step 2 and by definition of Γ.

Step 5: Non-anticipativity. With the same notation as in Step 4, on A, we have that for all $t\geq0$,

$$\begin{aligned}\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W) &= \lim_n \Gamma_t^n(\xi,X,\xi',Y,\xi^\dagger,W)\\ &= \lim_n \Gamma_t^n(\xi,(X_{s\wedge t})_{s\geq0},\xi',Y,\xi^\dagger,W)\\ &= \Gamma_t(\xi,(X_{s\wedge t})_{s\geq0},\xi',Y,\xi^\dagger,W)\end{aligned}$$

by (22). The last equality is legitimate because $(X_{s\wedge t})_{s\geq0}$ satisfies the conditions of Steps 2 and 3. This implies (5). □

We now turn to Lemma 6, using the non-anticipativity of Γ from Lemma 5 and the domination estimate from Proposition 11.

Proof of Lemma 6. Fix an admissible probability measure μ and an admissible pair (ξ,X) defined on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq0},\mathbb{P})$. Let $(\Omega^*,\mathcal{F}^*,(\mathcal{F}^*_t)_{t\geq0},\mathbb{P}^*)$ carry an admissible pair (ξ',Y) with law μ, let $(\Omega^\dagger,\mathcal{F}^\dagger,\mathbb{P}^\dagger)$ carry an $E^\dagger$-valued random variable $\xi^\dagger$ with law $\pi^\dagger$, and let $(\Omega^\#,\mathcal{F}^\#,(\mathcal{F}^\#_t)_{t\geq0},\mathbb{P}^\#)$ carry a one-dimensional Brownian motion W. These auxiliary spaces are taken independent of $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq0},\mathbb{P})$ and of each other. As in the proof of Lemma 12, we have

$$\Lambda_\mu(t,\xi,X) = \mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\phi(\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W))\gamma(X_t,Y_t)\Big]\Big]\Big].$$

For all $T>0$, there is a constant C_T such that a.s.,

$$\Psi_{\mu,T}(\xi,X) := \mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\sup_{t\in[0,T]}|\phi(\Gamma_t(\xi,X,\xi',Y,\xi^\dagger,W))|\Big]\Big]\Big] \leq C_T. \tag{25}$$

If ϕ is bounded, this is obvious. Else, this follows from Proposition 11-(i), the fact that ϕ has at most linear growth and (6).

Point (i), i.e. continuity of $(\Lambda_\mu(t,\xi,X))_{t\geq0}$, follows from dominated convergence by (25), the continuity and boundedness of γ, the continuity of ϕ and the facts that Γ is valued in $\mathcal{C}_1$ while X and Y are valued in $\mathcal{C}_d$.

For (ii), i.e. $(\mathcal{F}_t)_{t\geq0}$-adaptedness of $(\Lambda_\mu(t,\xi,X))_{t\geq0}$, it suffices to use (5): we have

$$\Lambda_\mu(t,\xi,X) = \mathbb{E}^*\Big[\mathbb{E}^\dagger\Big[\mathbb{E}^\#\Big[\phi(\Gamma_t(\xi,(X_{s\wedge t})_{s\geq0},\xi',Y,\xi^\dagger,W))\gamma(X_t,Y_t)\Big]\Big]\Big].$$

which is of course $\mathcal{F}_t$-measurable. □

References

[1] David J. Aldous. Representations for partially exchangeable arrays of random variables. Journal of Multivariate Analysis, 11(4):581–598, 1981.

[2] Siva Athreya, Frank Den Hollander, and Adrian Röllin. Graphon-valued stochastic processes from population genetics. The Annals of Applied Probability, 31(4):1724–1745, 2021.

[3] Erhan Bayraktar, Suman Chakraborty, and Ruoyu Wu. Graphon mean field systems. The Annals of Applied Probability, 33(5):3587–3619, 2023.

[4] Erhan Bayraktar and Ruoyu Wu. Mean field interaction on random graphs with dynamically changing multi-color edges. Stochastic Processes and their Applications, 141:197–244, 2021.

[5] Guo-qiang Bi and Mu-ming Poo. Synaptic modification by correlated activity: Hebb's postulate revisited. Annual Review of Neuroscience, 24:139–166, 2001.

[6] Peter Braunsteins, Frank Den Hollander, and Michel Mandjes. A sample-path large deviation principle for dynamic Erdős–Rényi random graphs. The Annals of Applied Probability, 33(4):3278–3320, 2023.

[7] Peter Braunsteins, Frank den Hollander, and Michel Mandjes. Graphon-valued processes with vertex-level fluctuations. Stochastic Processes and their Applications, 198:Paper No. 104961, 2026.

[8] Jiří Černỳ and Anton Klimovsky. Markovian dynamics of exchangeable arrays. In Genealogies of Interacting Particle Systems, pages 209–228. World Scientific, 2020.

[9] Hayato Chiba and Georgi S. Medvedev. The mean field analysis of the Kuramoto model on graphs I. the mean field equation and transition point formulas. Discrete and Continuous Dynamical Systems, 39(1):131–155, 2018.

[10] Harry Crane. Dynamic random networks and their graph limits. The Annals of Applied Probability, 26(2):691–721, 2016.

[11] Harry Crane. Exchangeable graph-valued Feller processes. Probability Theory and Related Fields, 168(3):849–899, 2017.

[12] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. Probability Theory and Related Fields, 162(3–4):707–738, 2015.

[13] Ankan Ganguly, Konstantinos Spiliopoulos, and Daniel Sussman. Mean-field analysis of latent variable process models on dynamically evolving graphs with feedback effects. arXiv preprint arXiv:2502.04280, 2025.

[14] Marios A Gkogkas, Christian Kuehn, and Chuang Xu. Continuum limits for adaptive network dynamics. Communications in Mathematical Sciences, 21(1):83–106, 2023.

[15] Marios Antonios Gkogkas, Christian Kuehn, and Chuang Xu. Mean field limits of co-evolutionary signed heterogeneous networks. European Journal of Applied Mathematics, pages 1–44, 2025.

[16] Thilo Gross and Bernd Blasius. Adaptive coevolutionary networks: a review. Journal of the Royal Society Interface, 5(20):259–271, 2008.

[17] Douglas N. Hoover. Relations on probability spaces and arrays of random variables. Preprint, Institute for Advanced Study, 1979.

[18] Pierre-Emmanuel Jabin, David Poyato, and Juan Soler. Mean-field limit of non-exchangeable systems. Communications on Pure and Applied Mathematics, 78(4):651–741, 2025.

[19] Pierre-Emmanuel Jabin, Valentin Schmutz, and Datong Zhou. Non-exchangeable networks of integrate-and-fire neurons: spatially-extended mean-field limit of the empirical measure. arXiv preprint arXiv:2409.06325, 2024.

[20] Pierre-Emmanuel Jabin and Datong Zhou. The mean-field limit of sparse networks of integrate-and-fire neurons. Annales de l'Institut Henri Poincaré C. Analyse Non Linéaire, 43(2):273–343, 2026.

[21] Svante Janson and Persi Diaconis. Graph limits and exchangeable random graphs. Rendiconti di Matematica e delle sue Applicazioni. Serie VII, pages 33–61, 2008.

[22] Mark Kac. Foundations of kinetic theory. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, volume 3, pages 171–197, Berkeley and Los Angeles, California, 1956. University of California Press.

[23] Dmitry Kaliuzhnyi-Verbovetskyi and Georgi S. Medvedev. The mean field equation for the Kuramoto model on graph sequences with non-Lipschitz limit. SIAM Journal on Mathematical Analysis, 50(3):2441–2465, 2018.

[24] Olav Kallenberg. On the representation theorem for exchangeable arrays. Journal of Multivariate Analysis, 30(1):137–154, 1989.

[25] Olav Kallenberg. Probabilistic Symmetries and Invariance Principles. Springer Science & Business Media, 2005.

[26] Rajeeva L. Karandikar. On pathwise stochastic integration. Stochastic Processes and their Applications, 57(1):11–18, 1995.

[27] Henry P. McKean. Propagation of chaos for a class of non-linear parabolic equations. In A. K. Aziz, editor, Lecture Series in Differential Equations, volume 2 of Van Nostrand Mathematical Studies, pages 177–194. Van Nostrand Reinhold Company, 1969.

[28] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), volume 1627 of Lecture Notes in Math., pages 42–95. Springer, Berlin, 1996.

[29] Daniel Revuz and Marc Yor. Continuous Martingales and Brownian Motion, volume 293 of Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, third edition, 1999.

[30] Alain-Sol Sznitman. Topics in propagation of chaos. Ecole d'été de probabilités de Saint-Flour XIX—1989, 1464:165–251, 1991.

[31] Datong Zhou. Non-exchangeable mean-field theory for adaptive weights: propagation of dissociatedness and graphon sampling lemma. arXiv preprint arXiv:2506.13587, 2025.

Nicolas Fournier, Sorbonne Université - LPSM (UMR 8001), Campus Pierre et Marie Curie, Case courrier 158, 4 place Jussieu, 75252 Paris Cedex 05, France. nicolas.fournier@sorbonne-universite.fr

Datong Zhou, Sorbonne Université - LPSM (UMR 8001) and LJLL (UMR 7598), Campus Pierre et Marie Curie, Case courrier 158, 4 place Jussieu, 75252 Paris Cedex 05, France. datong.zhou@sorbonne-universite.fr