Asymptotics of Interacting Stochastic Processes on Sparse Graphs

Kavita Ramanan
Division of Applied Math, Brown University

CRM PIMS Summer School in Probability 2021: MiniCourse 2
June 2021

Prelude to the Course

Problem Description

Networks of interacting stochastic processes

Given a finite connected graph $G = (V, E)$,

write $u \sim v$ if $(u, v) \in E$,

$N_v = N_v(G) = \{u \in V : u \sim v\}$ denotes the neighborhood of v,

$d_v = d_v(G) = |N_v(G)|$ denotes the degree of vertex v

Each node $v \in V$ has a particle whose stochastic evolution depends only on its own state and (symmetrically) on its neighbors' states

Networks of interacting stochastic processes

Given a finite connected graph $G = (V, E)$,

write $u \sim v$ if $(u, v) \in E$,

$N_v = N_v(G) = \{u \in V : u \sim v\}$ denotes the neighborhood of v,

$d_v = d_v(G) = |N_v(G)|$ denotes the degree of vertex v

Each node $v \in V$ has a particle whose stochastic evolution depends only on its own state and (symmetrically) on its neighbors' states

In this course, we will focus on two types of dynamics:

A. Discrete-time (Markov) processes

B. Diffusions

Such models describe a range of phenonena in statistical physics, epidemiology, neuroscience, math finance, etc.

Some basic notation

- A space E is said to be Polish if it is a complete, separable metrisable space
- for any Polish space E, let $\mathcal{P}(E)$ denote the space of Borel probability measures equipped with the topology of weak convergence: $\nu_n \to \nu$ in $\mathcal{P}(E)$ if for every bounded continuous $h : E \to \mathbb{R}$,

$$\int_E h(x)\nu_n(dx) \to \int_E h(x)\nu(dx).$$

- Then $\mathcal{P}(E)$ is also a Polish space.
- Also, given a Polish space E and $x \in E$, $\delta_x \in \mathcal{P}(E)$ denotes the Dirac delta measure: for any Borel set $A \subset E$, $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise.

A. Networks of interacting Markov chains

- Fix graph $G = (V, E)$ and initial condition $x = (x_v)_{v \in V} \in \mathcal{X}^V$

discrete-time Markov chain: for $v \in V$,

$$\mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{t}+1) = \mathbf{F}\left(\mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{t}), \mathbf{X}_{\mathbf{N_v}}^{\mathbf{G,x}}(\mathbf{t}), \xi_{\mathbf{v}}(\mathbf{t}+1)\right), \qquad \mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{0}) = x_v$$

where $X_A := (X_v)_{v \in A}$, in particular $X_{N_v}(t) = (X_u(t))_{u \sim v}$, and

A. Networks of interacting Markov chains

- Fix graph $G = (V, E)$ and initial condition $x = (x_v)_{v \in V} \in \mathcal{X}^V$

discrete-time Markov chain: for $v \in V$,

$$\mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{t}+\mathbf{1}) = \mathbf{F}\left(\mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{t}), \mathbf{X}_{\mathbf{N_v}}^{\mathbf{G,x}}(\mathbf{t}), \xi_{\mathbf{v}}(\mathbf{t}+\mathbf{1})\right), \qquad \mathbf{X}_{\mathbf{v}}^{\mathbf{G,x}}(\mathbf{0}) = x_v$$

where $X_A := (X_v)_{v \in A}$, in particular $X_{N_v}(t) = (X_u(t))_{u \sim v}$, and

- the state space $\mathcal{X}$ and noise space Σ are Polish
- $\xi_v(t), v \in V,\ t = 0, 1, \ldots,$ are i.i.d. Σ-valued noises
- continuous transition function $F : \mathcal{X} \times S^{\sqcup}(\mathcal{X}) \times \Sigma \to \mathcal{X}$ where $S^{\sqcup} = \bigsqcup_k S^k(\mathcal{X})$ is the disjoint union of (unordered) sequences of length k in $\mathcal{X}$: $S^k(\mathcal{X}) = \mathcal{X}^k/\mathrm{Sym}_k$ and Sym_k is the group of permutations on $[k]$.
- continuity in the sense that on inputs of length k, $F(x, \cdot, \xi) = F^k$ for some continuous $F^k : \mathcal{X} \times \mathcal{X}^k/\mathrm{Sym}^k$.

A comment on the transition function F

Probabilistic cellular automata, synchronous Markov chains

$$\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}+\mathbf{1}) = \mathbf{F}\left(\mathbf{X}_{\mathbf{v}}(\mathbf{t}), \mathbf{X}_{\mathbf{N}_{\mathbf{v}}}(\mathbf{t}), \xi_{\mathbf{v}}(\mathbf{t}+\mathbf{1})\right), \qquad \mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{0}) = x_v$$

where $F : \mathcal{X} \times S^{\sqcup}(\mathcal{X}) \times \Sigma \to \mathcal{X}$ is continuous

- A generic example is when F depends on the local empirical measure of the neighborhood of v:
  $\mu_v^{G,x}(t) = \frac{1}{d_v} \sum_{u \sim v} \delta_{X_u^{G,x}(t)}$, and

  $$F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)) = \bar{F}\left(X_v(t), \mu_v^{G,x}(t), \xi_v(t+1)\right)$$

  for some continuous $\bar{F} : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Sigma \mapsto \mathcal{X}$.
- But (i) $\bar{F}$ cannot distinguish between configurations $\{0,1\}$ and $\{0,0,1,1\}$ so cannot account for the number of occurences. (ii) requiring $\bar{F}$ to be continuous wrt $\mathcal{P}(\mathcal{X})$ may be too restrictive (e.g., may rule out a function of $\max_{u \sim v} x_u$)

Example: Discrete-time Contact Process

- State space $\mathcal{X} = \{0, 1\} = \{\text{healthy}, \text{infected}\}$.
- Parameters $p, q \in [0, 1]$.
- $X_v(t) \in \mathcal{X}$, state of particle at v at time t

Transition rule: At time t, evolution of state of particle at any (non-isolated) node v depends on state of particle at any v and the neighbors' empirical distribution at that time:

$$\mu_v(t) = \frac{1}{d_v} \sum_{u \sim v} \delta_{X_u(t)}$$

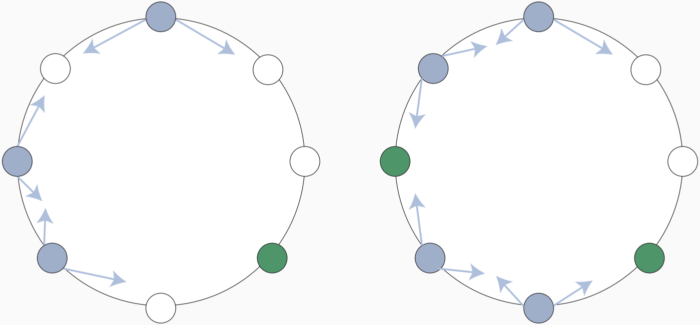
- if state $X_v(t) = 1$, it switches to $X_v(t+1) = 0$ w.p. q,
- if state $X_v(t) = 0$, it switches to $X_v(t+1) = 1$ w.p.

$$\frac{p}{d_v} \sum_{u \sim v} X_u(t) = p \int y \mu_v(t)(dy)$$

where recall d_v = degree of vertex v.

Example: Susceptible-Infected-Recovered (SIR) Process

$$S \xrightarrow{p\times(\text{frac. infected neighbors})} I \xrightarrow{q} R$$



B. Networks of interacting diffusions

- Fix a finite graph $G = (V, E)$
- initial condition $x = (x_v)_{v \in V} \in \mathbb{R}^{d^V}$ for some $d \in \mathbb{N}$

Evolves as a diffusion:

$$dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)$$

with $X_v^{G,x}(0) = x_v$, where

- drift coefficient $b : \mathbb{R}^d \times S^{\sqcup}(\mathbb{R}^d) \mapsto \mathbb{R}^d$ Lip. cont.
- diffusion coefficient $\sigma : \mathbb{R}^d \times S^{\sqcup}(\mathbb{R}^d) \mapsto \mathbb{R}^{d \times d}$ Lip. cont.
- i.i.d. d-dimensional Brownian motions $W_v, v \in V$.

B. Networks of interacting diffusions

- Fix a finite graph $G = (V, E)$
- initial condition $x = (x_v)_{v \in V} \in {\mathbb{R}^d}^V$ for some $d \in \mathbb{N}$

Evolves as a diffusion:

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)}$$

with $\mathbf{X_v^{G,x}(0) = x_v}$, where

- drift coefficient $b : \mathbb{R}^d \times S^{\sqcup}(\mathbb{R}^d) \mapsto \mathbb{R}^d$ Lip. cont.
- diffusion coefficient $\sigma : \mathbb{R}^d \times S^{\sqcup}(\mathbb{R}^d) \mapsto \mathbb{R}^{d \times d}$ Lip. cont.
- i.i.d. d-dimensional Brownian motions $W_v, v \in V$.
- Note each $\mathbf{X_v^{G,x}}$ takes values in $\mathcal{C}_d := \mathcal{C}([0, \infty) : \mathbb{R}^d)$,

Remark: Can consider more general, non-Markovian SDEs, with time-dependent and progressively measurable coefficients $b(t, X_v^{G,x}, X_{N_v(G)}^{G,x})$, but we will restrict to the above for simplicity

Example: Systemic Risk

Given independent Brownian motions $W_v, v = 1, \ldots, n$,

$$\mathbf{dX_v(t) = -hU(X_v(t))dt} + \theta(\mathbf{\overline{X}_v(t) - X_v(t))dt} + \sigma \mathbf{dW_v(t)},$$

for some restoring potential $U : \mathbb{R} \mapsto \mathbb{R}$, $\theta, \sigma > 0, h \in \mathbb{R}$, and with some given initial conditions, where $\overline{X}_v(t)$ is the local empirical mean:

$$\overline{X}_v(t) := \frac{1}{d_v} \sum_{u \sim v} X_u(t) = \int y \mu_v(t)(dy), \quad \mu_v(t) = \frac{1}{d_v} \sum_{u \sim v} \delta_{X_u(t)}$$

- $X_v(t)$ represents the state of risk of agent/component v
- Systemic risk is the risk that in an interconnected system of agents that can fail individually, a large number of them fails simultaneously, or nearly so.
- The interconnectivity of the agents and the form of evolution, play an essential role in systemic risk assessment.
- Is most well understood when $G = K_n$, the complete graph

Global empirical measures

Fix $G = (V, E)$. For $v \in V$, with $X_v^{G,x}(0) = x_v$, and

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)},$$

or

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)}.$$

Quantities of interest include

- the (global) empirical measure on path space

$$\mu^{G,x} := \frac{1}{|G|} \sum_{v \in G} \delta_{X_v^{G,x}}$$

Note that $\mu^{G,x}$ is a random element of $\mathcal{P}(\mathcal{X}^\infty)$ or $\mathcal{P}(\mathcal{C}_d)$;

- and the (global) empirical measure process

$$\mu^{G,x}(t) := \frac{1}{|G|} \sum_{v \in G} \delta_{X_v^{G,x}(t)}$$

Note that each $\mu^{G,x}(t)$ is a random element of $\mathcal{P}(\mathcal{X})$ or $\mathcal{P}(\mathbb{R}_d)$,

Summary: Networks of Interacting Stochastic Processes

$$X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right),$$

$$dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t).$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}(t)}$$

Summary: Networks of Interacting Stochastic Processes

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)},$$

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)}.$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}(t)}$$

Key questions: Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X^{G_n,x^n}}$ converge in a suitable sense?

Summary: Networks of Interacting Stochastic Processes

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)},$$

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)}.$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}(t)}$$

Key questions: Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X^{G_n,x^n}}$ converge in a suitable sense?

Q2. Do the global empirical measures μ^{G_n,x^n} converge?

Summary: Networks of Interacting Stochastic Processes

$$\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}+1) = \mathbf{F}\left(\mathbf{X}_{\mathbf{v}}(\mathbf{t}), (\mathbf{X}_{\mathbf{N}_{\mathbf{v}}}(\mathbf{t})), \xi_{\mathbf{v}}(\mathbf{t}+1)\right),$$

$$\mathbf{dX}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}) = \mathbf{b}(\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}), \mathbf{X}_{\mathbf{N}_{\mathbf{v}}(\mathbf{G})}^{\mathbf{G},\mathbf{x}}(\mathbf{t}))\mathbf{dt} + \sigma(\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}), \mathbf{X}_{\mathbf{N}_{\mathbf{v}}(\mathbf{G})}^{\mathbf{G},\mathbf{x}}(\mathbf{t}))\mathbf{dW}_{\mathbf{v}}(\mathbf{t}).$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v \in G} \delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v \in G} \delta_{X_v^{G,x}(t)}$$

Key questions: Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X}^{\mathbf{G_n},\mathbf{x^n}}$ converge in a suitable sense?

Q2. Do the global empirical measures μ^{G_n,x^n} converge?

Q3. can one autonomously characterize the limiting dynamics of a fixed or "typical particle" $\mathbf{X}_{\mathbf{v}}^{\mathbf{G_n},\mathbf{x_n}}(\mathbf{t}), t \in [0, T]$?

Background and Motivation: Classical Results

Classical Results in a Special Setting

For notational convenience, suppose $\sigma = I_d$ and the drift depends on neighbors only via their local empirical measure:

$$\mathbf{dX_v^{G_n,x^n}(t) = B(X_v^{G_n,x^n}(t), \mu_v^{G_n,x^n}(t))dt + dW_v(t)},$$

where recall $\mu_v^{G_n,x^n}(t)$ is the local empirical measure at t:

$$\mu_v^{G_n,x^n}(t) = \frac{1}{d_v} \sum_{u \in N_v} \delta_{X_u^{G_n,x^n}(t)}$$

for some continuous $B : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$, e.g. the linear case:

$$B(x, m) := \int_{\mathbb{R}^d} \overline{b}(x, y) m(dy), \qquad \overline{b} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$$

this reduces to the following SDE with "pairwise interactions"

$$\mathbf{dX_v^{G_n,x^n}(t) = \frac{1}{d_v(G_n)} \sum_{u \in N_v} \overline{b}(X_v^{G_n,x^n}(t), X_u^{G_n,x^n}(t))dt + dW_v(t)},$$

A well-studied case: when $G_n = K_n$ the complete graph

In this case, $V_n = [n]$, by setting $X^n = X^{K_n, x_n}$ and slightly modifying B, we can write

$$\mathbf{dX^n_v(t) = B(X^n_v(t), \mu^n(t))dt + dW_v(t)}, \quad v \in [n],$$

where $\mu^n(t) = \frac{1}{n}\sum_{v=1}^n \delta_{X_v^n(t)}$ is the global empirical measure.

Gaining intuition about the structure of the limit

- If $B = 0$, and $(X_v^n(0)), n \in \mathbb{N}$, are i.i.d. with law λ, independent of (W_v), (X_v^n) are i.i.d. Brownian motions. So the SLLN implies that a.s. μ^n converges weakly to the deterministic measure μ equal to $\mathrm{Law}(X_\rho(0) + W)$, $X_\rho(0) \sim \lambda$

A well-studied case: when $G_n = K_n$ the complete graph

In this case, $V_n = [n]$, by setting $X^n = X^{K_n, x_n}$ and slightly modifying B, we can write

$$\mathbf{dX_v^n(t) = B(X_v^n(t), \mu^n(t))dt + dW_v(t)}, \quad v \in [n],$$

where $\mu^n(t) = \frac{1}{n}\sum_{v=1}^n \delta_{X_v^n(t)}$ is the global empirical measure.

Gaining intuition about the structure of the limit

- If $B = 0$, and $(X_v^n(0)), n \in \mathbb{N}$, are i.i.d. with law λ, independent of (W_v), (X_v^n) are i.i.d. Brownian motions. So the SLLN implies that a.s. μ^n converges weakly to the deterministic measure μ equal to $\mathrm{Law}(X_\rho(0) + W)$, $X_\rho(0) \sim \lambda$
- In general if μ^n converges to a deterministic measure-valued process μ then since B is continuous, one might expect each $X_v^n \Rightarrow X_\rho$, where $X_\rho(0) \sim \lambda$ (ρ denotes a typical vertex) and

$$\mathbf{dX_\rho(t) = B(X_\rho(t), \mu(t))dt + dW_v(t)},$$

A well-studied case: when $G_n = K_n$ the complete graph

In this case, $V_n = [n]$, by setting $X^n = X^{K_n, x_n}$ and slightly modifying B, we can write

$$\mathbf{dX^n_v(t) = B(X^n_v(t), \mu^n(t))dt + dW_v(t)}, \quad v \in [n],$$

where $\mu^n(t) = \frac{1}{n}\sum_{v=1}^n \delta_{X^n_v(t)}$ is the global empirical measure.

Gaining intuition about the structure of the limit

- If $B = 0$, and $(X^n_v(0)), n \in \mathbb{N}$, are i.i.d. with law λ, independent of (W_v), (X^n_v) are i.i.d. Brownian motions. So the SLLN implies that a.s. μ^n converges weakly to the deterministic measure μ equal to $\mathrm{Law}(X_\rho(0) + W)$, $X_\rho(0) \sim \lambda$
- In general if μ^n converges to a deterministic measure-valued process μ then since B is continuous, one might expect each $X^n_v \Rightarrow X_\rho$, where $X_\rho(0) \sim \lambda$ (ρ denotes a typical vertex) and

$$\mathbf{dX_\rho(t) = B(X_\rho(t), \mu(t))dt + dW_v(t)},$$

with $\mu(t) = \mathrm{Law}(\mathbf{X}_\rho(\mathbf{t}))$ or in fact, $\mu = \mathrm{Law}(\mathbf{X}_\rho)$

Mean-field limits and nonlinear processes: $G_n = K_n$

Theorem, McKean '67; Oelschläger '84; 'Sznitman '91, etc.

If $(\mathbf{X}^{\mathbf{n}}_{\mathbf{v}}(\mathbf{0}))$, $n \in \mathbb{N}$, are i.i.d. with common law λ and B is Lipschitz continuous, then $(\mu^n(t))_{t\in[0,T]}$ converges in probability to the unique solution $(\mu(t))_{t\in[0,T]}$ of the McKean-Vlasov equation

$$\mathbf{dX(t) = B(X(t)}, \mu(\mathbf{t})\mathbf{)dt + dW(t)}, \qquad \mu(t) = \mathrm{Law}(\mathbf{X(t)}).$$

with $\mathbf{X(0)} \sim \lambda$, independent of the driving Brownian motion W. Moreover, the particles become asymptotically independent. Precisely, for fixed k,

$$(X_1^n, \ldots, X_k^n) \Rightarrow \mu^{\otimes k}, \qquad \text{as } n \to \infty.$$

Mean-field limits and nonlinear processes: $G_n = K_n$

Theorem, McKean '67; Oelschläger '84; 'Sznitman '91, etc.

If $(\mathbf{X_v^n(0)})$, $n \in \mathbb{N}$, are i.i.d. with common law λ and B is Lipschitz continuous, then $(\mu^n(t))_{t\in[0,T]}$ converges in probability to the unique solution $(\mu(t))_{t\in[0,T]}$ of the McKean-Vlasov equation

$$\mathbf{dX(t) = B(X(t), \mu(t))dt + dW(t)}, \qquad \mu(t) = \mathrm{Law}(\mathbf{X(t)}).$$

with $\mathbf{X(0)} \sim \lambda$, independent of the driving Brownian motion W.
Moreover, the particles become asymptotically independent.
Precisely, for fixed k,

$$(X_1^n, \ldots, X_k^n) \Rightarrow \mu^{\otimes k}, \qquad \text{as } n \to \infty.$$

More generally the above results hold when $\mu^n(0) \Rightarrow \lambda$
The phenomenon is referred to as propagation of chaos.

Note: X is a Markov process

Mean-field limits and nonlinear processes: $G_n = K_n$

Theorem, McKean '67; Oelschläger '84; 'Sznitman '91, etc.

If $(\mathbf{X^n_v(0)})$, $n \in \mathbb{N}$, are i.i.d. with common law λ and B is Lipschitz continuous, then $(\mu^n(t))_{t\in[0,T]}$ converges in probability to the unique solution $(\mu(t))_{t\in[0,T]}$ of the McKean-Vlasov equation

$$\mathbf{dX(t) = B(X(t), \mu(t))dt + dW(t)}, \qquad \mu(t) = \mathrm{Law}(\mathbf{X(t)}).$$

with $\mathbf{X(0)} \sim \lambda$, independent of the driving Brownian motion W. Moreover, the particles become asymptotically independent. Precisely, for fixed k,

$$(X^n_1, \ldots, X^n_k) \Rightarrow \mu^{\otimes k}, \qquad \text{as } n \to \infty.$$

More generally the above results hold when $\mu^n(0) \Rightarrow \lambda$
The phenomenon is referred to as propagation of chaos.

Note: X is a Markov process with a nonlinear fwd Kolmogorov eqn.

A Slightly Different Perspective

When $G_n = K_n$, the existence of a limit for $X_v^n = X_v^{K_n, x^n}$ follows from general results on exchangeable processes:

Kurtz and Kotelenez ('10)

As $n \to \infty$, $X_v^n \Rightarrow X_v^\infty$, where

$$dX_v^\infty(t) = b(X_i^\infty(t), \mu(t))dt + dW_i(t), \qquad v = 1, 2, \ldots,$$

with the following limit existing:

$$\mu(t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i^n(t)}.$$

A Slightly Different Perspective

When $G_n = K_n$, the existence of a limit for $X_v^n = X_v^{K_n, x^n}$ follows from general results on exchangeable processes:

Kurtz and Kotelenez ('10)

As $n \to \infty$, $X_v^n \Rightarrow X_v^\infty$, where

$$dX_v^\infty(t) = b(X_i^\infty(t), \mu(t))dt + dW_i(t), \qquad v = 1, 2, \ldots,$$

with the following limit existing:

$$\mu(t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{X_i^n(t)}.$$

The McKean-Vlasov limit

ginal dynamics of an infinite particle system

Propagation of Chaos for Dense Graph Sequences

- If $G_n \neq K_n$ but the sequence (G_n) is dense, in the sense that the degrees in the graph G_n are diverging to infinity as $n \to \infty$, then $\mathbf{X_v^{G_n,x^n}}, v \in G_n$, are still weakly interacting: e.g.,

$$\mathbf{dX_v^{G_n,x^n}(t) = \frac{1}{d_v(G_n)} \sum_{u \in N_v} \bar{b}(X_v^{G_n,x^n}, X_u^{G_n,x^n})dt + dW_t}$$

Propagation of Chaos for Dense Graph Sequences

- If $G_n \neq K_n$ but the sequence (G_n) is dense, in the sense that the degrees in the graph G_n are diverging to infinity as $n \to \infty$, then $\mathbf{X_v^{G_n,x^n}}$, $v \in G_n$, are still weakly interacting: e.g.,

$$\mathbf{dX_v^{G_n,x^n}(t) = \frac{1}{d_v(G_n)} \sum_{u \in N_v} \bar{b}(X_v^{G_n,x^n}, X_u^{G_n,x^n})dt + dW_t}$$

- So still expect propagation of chaos and asymptotic independence, so that once again, the global empirical measure μ^{G_n,x^n} converges to a deterministic limit μ and so the typical particle dynamics converges again to the mean-field limit

$$\mathbf{dX(t) = \int_{\mathbb{R}^d} \bar{b}(X(t), y)\mu(t)(dy) + dW(t)}, \quad \mu(t) = \mathrm{Law}(\mathbf{X(t)}).$$

Propagation of Chaos for Dense Graph Sequences

- If $G_n \neq K_n$ but the sequence (G_n) is dense, in the sense that the degrees in the graph G_n are diverging to infinity as $n \to \infty$, then $\mathbf{X}_{\mathbf{v}}^{\mathbf{G_n},\mathbf{x^n}}$, $v \in G_n$, are still weakly interacting: e.g.,

$$\mathbf{dX_v^{G_n,x^n}(t) = \frac{1}{d_v(G_n)} \sum_{u \in N_v} \bar{b}(X_v^{G_n,x^n}, X_u^{G_n,x^n})dt + dW_t}$$

- So still expect propagation of chaos and asymptotic independence, so that once again, the global empirical measure μ^{G_n,x^n} converges to a deterministic limit μ and so the typical particle dynamics converges again to the mean-field limit

$$\mathbf{dX(t) = \int_{\mathbb{R}^d} \bar{b}(X(t), y)\mu(t)(dy) + dW(t)}, \quad \mu(t) = \mathrm{Law}(\mathbf{X(t)}).$$

- Topology does not matter!

Propagation of Chaos for Dense Graph Sequences

- If $G_n \neq K_n$ but the sequence (G_n) is dense, in the sense that the degrees in the graph G_n are diverging to infinity as $n \to \infty$, then $\mathbf{X_v^{G_n,x^n}}$, $v \in G_n$, are still weakly interacting: e.g.,

$$\mathbf{dX_v^{G_n,x^n}(t) = \frac{1}{d_v(G_n)} \sum_{u \in N_v} \bar{b}(X_v^{G_n,x^n}, X_u^{G_n,x^n})dt + dW_t}$$

- So still expect propagation of chaos and asymptotic independence, so that once again, the global empirical measure μ^{G_n,x^n} converges to a deterministic limit μ and so the typical particle dynamics converges again to the mean-field limit

$$\mathbf{dX(t)} = \int_{\mathbb{R}^{\mathbf{d}}} \mathbf{\bar{b}(X(t), y)}\mu\mathbf{(t)(dy) + dW(t)}, \quad \mu(t) = \mathrm{L}aw(\mathbf{X(t)}).$$

- Topology does not matter!
- Many recent results of this nature: Delattre-Giacomin-Luçon '16; Delarue '17, Coppini-Dietert-Giacomin '18, Reis-Oliveira '18, Bhamidi-Budhiraja-Wu '19, etc.

Analogous Mean-Field Limits hold in Discrete-time

discrete-time Markov chain: if for $v \in V$,

$$\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}+\mathbf{1}) = \bar{\mathbf{F}}\left(\mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}), \mu_{\mathbf{V}}^{\mathbf{G},\mathbf{x}}(\mathbf{t}), \xi_{\mathbf{v}}(\mathbf{t}+\mathbf{1})\right), \qquad \mathbf{X}_{\mathbf{v}}^{\mathbf{G},\mathbf{x}}(\mathbf{0}) = x_v$$

for continuous $\bar{F} : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Sigma \to \mathcal{X}$

Mean-field limit for discrete-time chains

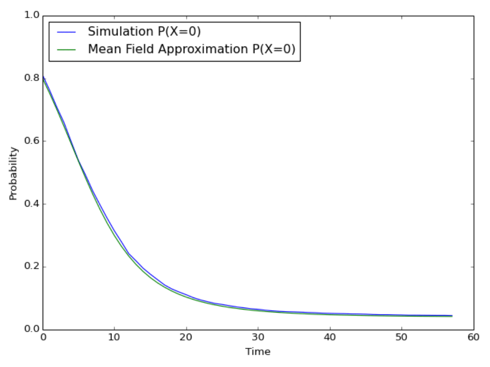
If $G_n = K_n$ and initial conditions are such that $\mu^{K_n,x^n}(0) \Rightarrow \mu(0)$, then the (random) global empirical measure sequence μ^{K_n,x^n} converges weakly to a deterministic measure sequence μ and for any v, $\mathbf{X}_{\mathbf{v}}^{\mathbf{K_n},\mathbf{x^n}}$ converges to the nonlinear discrete-time Markov chain $\mathbf{X}_\rho$:

$$\mathbf{X}_\rho(\mathbf{t}+\mathbf{1}) = \bar{F}(\mathbf{X}_\rho(\mathbf{t}), \mu_v(t), \xi_v), \quad \mu_v(t) = \mathrm{Law}(\mathbf{X}_{\mathbf{v}}(\mathbf{t})),$$

where $(\xi_v)_{v\in V}$ are i.i.d. noises with the same law as $\xi_v(1)$.

How well do mean-field approximations work?

Numerical results for the discrete-time SIR process on the complete graph

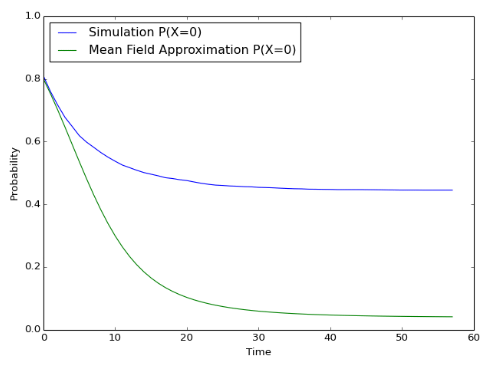


Plot of probability of being healthy vs. time
simulations due to Mitchell Wortsman

Mean-field approximation works well!!

How well do mean-field approximations work?

Numerical results for the discrete-time SIR Process
on the cycle graph



Plot of probability of being healthy vs. time
simulations due to Mitchell Wortsman

Mean-field approximation fails ...

Beyond Mean-Field Limits

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), X_{N_v}(t), \xi_v(t+1)\right)},$$

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}, X_{N_v(G)}^{G,x})dt + \sigma(X_v^{G,x}, X_{N_v(G)}^{G,x})dW_v(t)},$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}(t)}$$

Beyond Mean-Field Limits

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), X_{N_v}(t), \xi_v(t+1)\right)},$$

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}, X_{N_v(G)}^{G,x})dt + \sigma(X_v^{G,x}, X_{N_v(G)}^{G,x})dW_v(t)},$$

$$\mu^{G,x} := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}} \qquad \mu^{G,x}(t) := \frac{1}{|G|}\sum_{v\in G}\delta_{X_v^{G,x}(t)}$$

Our Focus is on Asymptotics for Sparse Graph Sequences

Example: G_n = Erdős-Rényi $\mathcal{G}(n, p_n)$ with $np_n \to p \in (0,\infty)$.
Open Question: Delattre-Giacomin-Luçon – to characterize typical dynamics

Outline of the Mini-Course

Outline

Recall Key questions: Given a sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X}^{\mathbf{G_n},\mathbf{x^n}}$ converge in a suitable sense?

Outline

Recall Key questions: Given a sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X}^{\mathbf{G_n},\mathbf{x^n}}$ converge in a suitable sense?

Q2. Do the global empirical measures μ^{G_n,x^n} converge?

Outline

Recall Key questions: Given a sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X}^{\mathbf{G_n},\mathbf{x^n}}$ converge in a suitable sense?

Q2. Do the global empirical measures μ^{G_n,x^n} converge?

Q3. can one autonomously characterize the limiting dynamics of a fixed or "typical particle" $P(\mathbf{X}_{\mathbf{v}}^{\mathbf{G_n},\mathbf{x_n}}(\mathbf{t}), t \in [0, T]$?

Outline

Recall Key questions: Given a sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, and appropriate initial conditions $x^n \in \mathcal{X}^{V_n}$,

Q1. Do the processes $\mathbf{X}^{\mathbf{G_n}, \mathbf{x^n}}$ converge in a suitable sense?

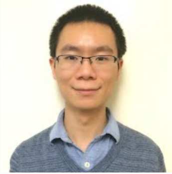
Q2. Do the global empirical measures μ^{G_n, x^n} converge?

Q3. can one autonomously characterize the limiting dynamics of a fixed or "typical particle" $P(\mathbf{X}_{\mathbf{v}}^{\mathbf{G_n}, \mathbf{x_n}}(\mathbf{t}), t \in [0, T]$?

- The rest of Lecture 1 will address Q1 (and p'haps part of Q2)
- Lecture 2 will address Q2 and discuss additional properties useful for Q3
- Lecture 3 will focus on Q3

Collaborators

Daniel Lacker
Columbia University



Ruoyu Wu
Iowa State University

Ankan Ganguly
Brown University

Mitchell Wortsman

Timothy-Sudijono

Mira Gordin

Most Directly Relevant References

- Oliveira, Reis, Stolerman, "Interacting diffusions on sparse graphs: hydrodynamics from local weak limits," *EJP* **25** (2020).
- Lacker, R., Wu, "Large sparse networks of interacting diffusions," Arxiv Preprint (2019)
- Lacker, R., Wu, "Local weak convergence for sparse networks of interacting processes," Arxiv Preprint (2020)
- Lacker, R., Wu, "Locally interacting diffusions as space-time Markov random fields," Arxiv Preprint (2019)
- Lacker, R., Wu, "Marginal dynamics of interacting diffusions on unimodular Galton-Watson trees," Arxiv Preprint (2020)
- Lacker, R., Wu, "Marginal Dynamics of probabilistic cellular automata on trees," Preprint (2021).
- Ganguly and R., "Limits of empirical measures of interacting particle systems on large sparse graphs," *near completion*, (2021)
- MacLaurin, "Large Deviations of a Network of Neurons with Dynamic Sparse Random Connections", Arxiv Preprint, 2016.

Rest of Lecture 1

Notion of local convergence

Process Convergence Result

Preliminaries on empirical measure convergence

Process Convergence Question

$$X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right),$$

$$dX_v^{G,x}(t) = b(X_v^{G,x}, X_{N_v(G)}^{G,x})dt + \sigma(X_v^{G,x}, X_{N_v(G)}^{G,x})dW_v(t),$$

Q1. Do the processes X^{G_n,x^n} converge in a suitable sense?

Process Convergence Question

$$X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right),$$

$$dX_v^{G,x}(t) = b(X_v^{G,x}, X_{N_v(G)}^{G,x})dt + \sigma(X_v^{G,x}, X_{N_v(G)}^{G,x})dW_v(t),$$

Q1. Do the processes X^{G_n,x^n} converge in a suitable sense?

- First step: for what sequences G_n can we expect this to hold?
- Taking inspiration form static models (Dembo-Montanari '10), when G_n converges locally to a limit graph G

Notion of local convergence

Local weak convergence of graphs

Idea: Encode sparsity via local weak convergence of graphs.
(a.k.a. Benjamini-Schramm convergence, see Aldous-Steele '03)

Local weak convergence of graphs

Idea: Encode sparsity via local weak convergence of graphs.
(a.k.a. Benjamini-Schramm convergence, see Aldous-Steele '03)

Definition: A graph $G = (V, E, \rho)$ is assumed to be rooted, finite or countable, locally finite, and connected.

Local weak convergence of graphs

Idea: Encode sparsity via local weak convergence of graphs.
(a.k.a. Benjamini-Schramm convergence, see Aldous-Steele '03)

Definition: A graph $G = (V, E, \rho)$ is assumed to be rooted, finite or countable, locally finite, and connected.

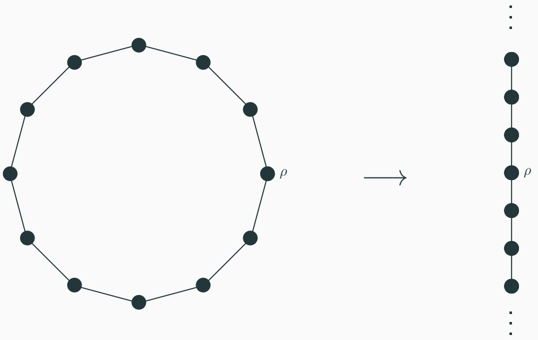
Definition: Rooted graphs G_n converge locally to G if:

$$\forall k \ \exists N \text{ s.t. } B_k(G) \cong B_k(G_n) \text{ for all } n \geq N,$$

where $B_k(\cdot)$ is ball of radius k at root, and $\cong$ means isomorphism.

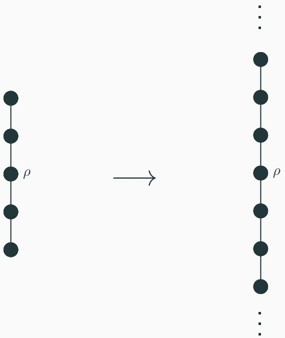
Examples of local weak convergence

1. **Cycle graph converges to infinite line**



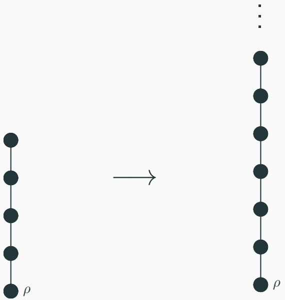
Examples of local weak convergence

2. **Line graph converges to infinite line**



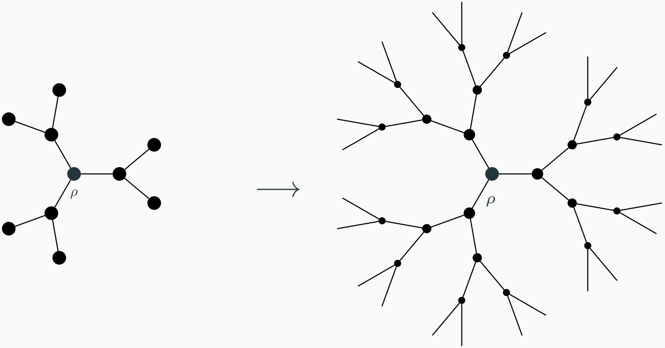
Examples of local weak convergence

3. **Line graph rooted at end converges to semi-infinite line**



Examples of local weak convergence

4. Finite to infinite d-regular trees
(A graph is d-regular if ever vertex has degree d.)

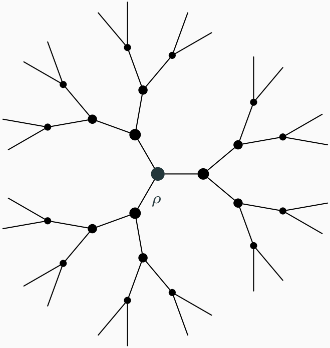


Examples of local weak convergence

5. Uniformly random regular graph to infinite regular tree

Fix d. Among all d-regular graphs on n vertices, select one uniformly at random. Place the root at a (uniformly) random vertex. When $n \to \infty$, this converges (in law) to the infinite d-regular tree. (McKay '81)

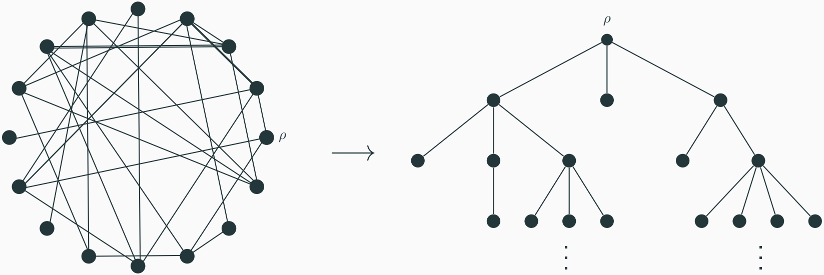
$\longrightarrow$



Examples of local weak convergence

6. Erdős-Rényi to Galton-Watson

If $G_n = G(n, p_n)$ with $np_n \to p \in (0, \infty)$, then G_n converges in law to the Galton-Watson tree with offspring distribution $\mathrm{Poisson}(p)$.



7. Preferential Attachment Graphs to a Random Tree
A result by Berger-Borgs-Chayes-Saberi ('14) shows convergence of preferential attachment graphs to a random tree

Key Point

Local limits of many classes of random graphs are often trees

Examples of Local weak convergence

7. Preferential Attachment Graphs to a Random Tree
A result by Berger-Borgs-Chayes-Saberi ('14) shows convergence of preferential attachment graphs to a random tree

Key Point

Local limits of many classes of random graphs are often trees

8. Convergence of Finite Lattices
$\mathbb{Z}^\kappa \cap [-n, n]^\kappa$ converges to $\mathbb{Z}^\kappa$

Local convergence of marked graphs

Recall: $G_n = (V_n, E_n, \rho_n)$ converges locally to $G = (V, E, \rho)$ if

$\forall k \; \exists N$ s.t. $B_k(G) \cong B_k(G_n)$ for all $n \geq N$.

Local convergence of marked graphs

Recall: $G_n = (V_n, E_n, \rho_n)$ converges locally to $G = (V, E, \rho)$ if

$$\forall k \; \exists N \text{ s.t. } B_k(G) \cong B_k(G_n) \text{ for all } n \geq N.$$

Definition: With G_n, G as above: Given a metric space (E, d_E) and a sequence $\boldsymbol{x}^n = (x_v^n)_{v \in G_n} \in E^{G_n}$, say that $(G_n, \boldsymbol{x}^n)$ converges locally to $(G, \boldsymbol{x})$ if

$$\forall k, \epsilon > 0 \; \exists N \text{ s.t. } \forall n \geq N \; \exists \varphi : B_k(G_n) \to B_k(G) \text{ isomorphism s.t. } \max_{v \in B_k(G_n)} d_E(x_v^n, x_{\varphi(v)}) < \epsilon.$$

Lemma

The set $\mathcal{G}_*[E]$ of (isomorphism classes of) $(G, \boldsymbol{x})$ admits a Polish topology compatible with the above convergence.

Process Convergence Result

Beyond Mean-Field Limits ...

$$\mathbf{X_v^{G,x}}(\mathbf{t}+\mathbf{1}) = \mathbf{F}\left(\mathbf{X_v}(\mathbf{t}), (\mathbf{X_{N_v}}(\mathbf{t})), \xi_{\mathbf{v}}(\mathbf{t}+\mathbf{1})\right),$$

with $(\xi_v(k))_{k\in\mathbb{N}, v\in G}$ i.i.d. with the same law for all graphs G;

$$\mathbf{dX_v^{G,x}(t)} = \mathbf{b}(\mathbf{X_v^{G,x}(t)}, \mathbf{X^{G,x}_{N_v(G)}(t)})\mathbf{dt} + \sigma(\mathbf{X_v^{G,x}(t)}, \mathbf{X^{G,x}_{N_v(G)}(t)})\mathbf{dW_v(t)},$$

with $\mathbf{X_v^{G,x}}(\mathbf{0}) = \mathbf{x_v}$, F continuous, b, σ Lipschitz continuous.

Let P^{G_n,x^n} be the law of marked graph $(G_n, \mathbf{X^n})$

Theorem 1: Lacker-Ramanan-Wu; '19/'20

1. For Markov chains, if $(G_n, x^n) \to (G, x)$ in $\mathcal{G}_*[\mathcal{X}]$ then $P^{G_n,x^n} \to P^{G,x}$ in $\mathcal{P}(\mathcal{G}_*[\mathcal{X}^\infty])$

Beyond Mean-Field Limits ...

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)},$$

with $(\xi_v(k))_{k\in\mathbb{N}, v\in G}$ i.i.d. with the same law for all graphs G;

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)},$$

with $\mathbf{X_v^{G,x}(0) = x_v}$, F continuous, b, σ Lipschitz continuous.

Let P^{G_n,x^n} be the law of marked graph $(G_n, \mathbf{X^n})$

Theorem 1: Lacker-Ramanan-Wu; '19/'20

1. For Markov chains, if $(G_n, x^n) \to (G, x)$ in $\mathcal{G}_*[\mathcal{X}]$ then $P^{G_n,x^n} \to P^{G,x}$ in $\mathcal{P}(\mathcal{G}_*[\mathcal{X}^\infty])$
2. For diffusions, if $\sup_{n\in\mathbb{N}} \sup_{v\in G_n} |x_v^n| \leq r$ and $(G_n, x^n) \to (G, x)$ in $\mathcal{G}_*[B_r(\mathbb{R}^d)]$, then $P^{G_n,x_n} \to P^{G,x}$ in $\mathcal{P}(\mathcal{G}_*[\mathcal{C}_d])$
3. Immediate extensions to the case of random graphs hold.

Beyond Mean-Field Limits ...

$$\mathbf{X_v^{G,x}(t+1) = F\left(X_v(t), (X_{N_v}(t)), \xi_v(t+1)\right)},$$

with $(\xi_v(k))_{k\in\mathbb{N},v\in G}$ i.i.d. with the same law for all graphs G;

$$\mathbf{dX_v^{G,x}(t) = b(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dt + \sigma(X_v^{G,x}(t), X_{N_v(G)}^{G,x}(t))dW_v(t)},$$

with $\mathbf{X_v^{G,x}(0) = x_v}$, F continuous, b, σ Lipschitz continuous.

Let P^{G_n,x^n} be the law of marked graph $(G_n, \mathbf{X^n})$

Theorem 1: Lacker-Ramanan-Wu; '19/'20

1. For Markov chains, if $(G_n, x^n) \to (G, x)$ in $\mathcal{G}_*[\mathcal{X}]$ then $P^{G_n,x^n} \to P^{G,x}$ in $\mathcal{P}(\mathcal{G}_*[\mathcal{X}^\infty])$
2. For diffusions, if $\sup_{n\in\mathbb{N}} \sup_{v\in G_n} |x_v^n| \le r$ and $(G_n, x^n) \to (G, x)$ in $\mathcal{G}_*[B_r(\mathbb{R}^d)]$, then $P^{G_n,x_n} \to P^{G,x}$ in $\mathcal{P}(\mathcal{G}_*[\mathcal{C}_d])$
3. Immediate extensions to the case of random graphs hold.

Next lecture:

Will discuss the proof of Theorem 1 and then move to study empirical measures