Scaling Limits
of
Interacting Particle Systems

Kavita Ramanan
Division of Applied Math
Brown University

11th Cornell Probability Summer School
Cornell, Ithaca
Jun 17–20, 2019
Lecture 1
Motivation and Introduction
1. Some Motivating Examples
Example 1: Spread of Infections

- Suppose you want to study how infections (virus in computers or diseases in the real world) spread in the real world.
- Some questions of interest:
  - How many times does a person/computer catch a particular disease/virus in a given time period?
  - What proportion of time is the person sick in a given period?
  - What proportion of time are more than half of the population sick?
Example 1: Discrete-time Contact Process

- Identify a graph (a collection of nodes, with edges between them representing connections) that describes the contact network of people in a population.
- Each node has two states: 1 if infected and 0 if sick.
- Estimate probability of catching an infection given the state of the population in your local contact neighborhood.

Various multi-state generalizations: SIS, SIR
Example 1: Discrete-time Contact Process

- State space $S = \{0, 1\} = \{\text{healthy, infected}\}$.
- Parameters $p, q \in [0, 1]$.
- $X_v(t) \in S$, state of process at time $t$
- The processes evolve as a discrete-time Markov chain

**Transition rule:** At time $t$, evolution of state of particle at node $v$ depends on state of particle at $v$ 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 $v$. 
Ising model

Probability distribution on $\{-1, 1\}^V$: for $\sigma \in \{-1, 1\}^V$,

$$
\mathbb{P}(X = \sigma) = \frac{1}{Z_{\beta}} e^{-\beta H(\sigma)},
$$

where $\beta > 0$ is a parameter referred to as the inverse temperature, and $H$ is the “Hamiltonian”:

$$
H(\sigma) = -J \sum_{i \sim j} \sigma_i \sigma_j - h \sum_j \sigma_j
$$

for suitable parameters $J, h \in \mathbb{R}$, and $Z_{\beta}$ is the normalizing constant

$$
Z_{\beta} = \sum_{\sigma \in \{-1,1\}^V} e^{-\beta H(\sigma)}.
$$
I. Ising model

Probability distribution on \{-1, 1\}^V: for \(\sigma \in \{-1, 1\}^V\),

\[
P(X = \sigma) = \frac{1}{Z_\beta} e^{-\beta H(\sigma)},
\]

where \(\beta > 0\) is a parameter referred to as the inverse temperature, and \(H\) is the “Hamiltonian”:

\[
H(\sigma) = -J \sum_{i \sim j} \sigma_i \sigma_j - h \sum_j \sigma_j
\]

for suitable parameters \(J, h \in \mathbb{R}\), and \(Z_\beta\) is the normalizing constant

\[
Z_\beta = \sum_{\sigma \in \{-1,1\}^V} e^{-\beta H(\sigma)}.
\]

Although there is an explicit expression for the probability, it is typically computationally infeasible to calculate \(Z_\beta\).
Instead, to (approximately) sample from the distribution

$$\mathbb{P}(X = \sigma) = \frac{1}{Z_\beta} e^{-\beta H(\sigma)},$$

one constructs a reversible Markov chain (the so-called Glauber dynamics) that has the target distribution as its stationary distribution.

**Note:** The transition probability matrix of the Markov chain only depends on ratios of the probabilities, and thus does not require knowledge of $Z_\beta$. 
Brownian Motion

Brownian motion \( \{W_t, t \geq 0\} \) is an \( \mathbb{R} \)-valued stochastic process such that

- \( t \mapsto W_t \) is (almost surely) continuous
- for every \( 0 < s < t < \infty \), \( W_t - W_s \) is independent of \( W_s \) and is distributed according to \( \mathcal{N}(0, t - s) \).
- \( W_0 = 0 \) (standard Brownian motion)
### Brownian Motion

Brownian motion \( \{ W_t, t \geq 0 \} \) is an \( \mathbb{R} \)-valued stochastic process such that

- \( t \mapsto W_t \) is (almost surely) continuous
- for every \( 0 < s < t < \infty \), \( W_t - W_s \) is independent of \( W_s \) and is distributed according to \( \mathcal{N}(0, t - s) \).
- \( W_0 = 0 \) (standard Brownian motion)

\( d \)-dimensional Brownian motion \( \{ B_t, t \geq 0 \} \) is an \( \mathbb{R}^d \)-valued stochastic process such that its components \( B^i, i = 1, \ldots, d \), are independent and identically distributed standard Brownian motions.
1. Some Motivating Examples

Example 3: Systemic Risk

- 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 its form of evolution, plays an essential role in systemic risk assessment.

\( X_t^i \) represents the state of risk of agent/component \( j \)

Given independent Brownian motions \( W^j, j = 1, \ldots, n \),

\[
dX_t^j = -hU(X_t^j)dt + \theta(\bar{X}_t - X_t^j)dt + \sigma dW_t^j,
\]

for some restoring potential \( U : \mathbb{R} \to \mathbb{R}, \theta, \sigma > 0 \), and with some given initial conditions, where \( \bar{X} \) is the empirical mean:

\[
\bar{X}_t := \frac{1}{n} \sum_{i=1}^{n} X_t^i.
\]
2. General Problem Formulation

Summary of Graph Terminology

- An (undirected) graph \( G = (V, E) \) consists of a countable collection \( V \) of vertices/nodes and a set \( E \) of edges, or (unordered) pairs in \( V \times V \).
- The neighboring relation is often denoted \( u \sim v \) if \((u, v) \in E\).
- A graph \( G \) is said to be finite if \(|V| < \infty\).
- The degree \( d_v \) of a vertex \( v \) is the cardinality of its neighborhood:
  \[
  d_v = |N_v| := |\{u \in V : u \sim v\}|
  \]
- An infinite graph \( G \) is said to be locally finite if \( d_v < \infty \) for every \( v \in V \).
- A graph \( G \) is said to be simple if it contains no loops (i.e., for every \( v \in V \), \((v, v) \notin E\)).
- The graph distance \( d(u, v) = d_G(u, v) \) between two vertices \( u \) and \( v \) is the length of the shortest path between the vertices.
Given a finite connected graph $G = (V, E)$, we are interested in a stochastic process

$$\{X_t^v, v \in G, t \in \mathbb{T}\},$$

with $\mathbb{T} = \mathbb{N}_0$ or $\mathbb{T} = [0, \infty)$ and the property that:

the evolution of the stochastic process is such that the state $X_t^v$ of each node $v \in V$ at time $t$ evolves stochastically depending only on its own state and those of its neighbors at that time.
For example, evolving as a \textit{discrete-time Markov chain}:

$$X_{t+1}^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_{t+1}^v \right), \ v \in V,$$

for a suitable function

$$F : S \times S^J \times U \mapsto S.$$

- state space $S$
- continuous transition function $F$
- independent noises $\xi_t^v, \ v \in V, \ t = 0, 1, \ldots$, taking values in some space $U$. 
For example, evolving as a discrete-time Markov chain:

\[ X_{t+1}^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_{t+1}^v \right), \quad v \in V, \]

for a suitable function

\[ F : S \times S^J \times U \mapsto S. \]

- state space \( S \)
- continuous transition function \( F \)
- independent noises \( \xi_t^v, v \in V, t = 0, 1, \ldots, \) taking values in some space \( U \).

Probabilistic cellular automata, synchronous Markov chains
Symmetric Dependence on Neighbors

The state of each \( v \in V \) evolves stochastically depending only on its own state and the empirical distribution of its neighbors.
Symmetric Dependence on Neighbors

The state of each $v \in V$ evolves stochastically depending only on its own state and the empirical distribution of its neighbors.

Then the discrete-time Markov chain takes the form:

$$X_{t+1}^v = \bar{F} \left( (X_t^v), \mu_t^v, \xi_{t+1}^v \right), \quad v \in V,$$

where

$$\bar{F} : S \times \mathcal{P}(S) \times U \mapsto S$$

and $\mu_t^v \in \mathcal{P}(S)$ is the local empirical measure at $v$:

$$\mu_t^v = \frac{1}{d_v} \sum_{u \sim v} \delta_{X_t^u}.$$
Brownian Motion

Brownian motion \( \{ W_t, \mathcal{F}_t, t \geq 0 \} \) on a filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})\) is an \( \mathbb{R} \)-valued stochastic process such that \( W_t \) is \( \mathcal{F}_t \)-measurable for every \( t \geq 0 \),

- \( t \mapsto W_t \) is (almost surely) continuous
- for every \( 0 < s < t < \infty \), \( W_t - W_s \) is independent of \( \mathcal{F}_s \) and is distributed according to \( \mathcal{N}(0, t - s) \).
Or as a diffusion:

\[ dX_t^v = \frac{1}{d_v} \sum_{u \sim v} b(X_t^v, X_t^u) dt + dW_t^v, \]

where \((W^v)^v \in V\) are independent \(d\)-dimensional Brownian motions.

For concreteness, assume standard conditions:
- \(b\) is Lipschitz and has linear growth, \(\sigma\) is non-degenerate and has an inverse,
- initial states \((X_0^v)^v \in V\) are i.i.d. and square-integrable.
We will focus on
discrete-time Markov chains and diffusions

but one can consider different interacting stochastic evolutions,
including
continuous time Markov chains (A. Ganguly),
jump diffusions, etc. ...
Key questions

Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, how can we describe the limiting behavior of...

- the dynamics of a fixed or “typical particle” $X^v_t$, $t \in [0, T]$?
- the empirical distribution of particles $\frac{1}{|V_n|} \sum_{v \in V_n} \delta X^v_t$?
Networks of interacting stochastic processes

Key questions

Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, how can we describe the limiting behavior of...

- the dynamics of a fixed or “typical particle” $X^v_t, t \in [0, T]$?
- the empirical distribution of particles $\frac{1}{|V_n|} \sum_{v \in V_n} \delta X^v_t$?

A much-studied special case

$G_n = K_n$, the complete graph on $n$ vertices
3. The mean field case (McKean-Vlasov 1966)

\( G_n = K_n \) complete graph; wlog \( V_n = \{1, \ldots, n\} \)

Particles \( i = 1, \ldots, n \) interact according to

\[
dX^i_t = \frac{1}{n} \sum_{k=1}^{n} b(X^i_t, X^k_t) dt + dW^i_t, \quad i = 1, \ldots, n
\]

where \( W^1, \ldots, W^n \) are independent Brownian motions, with iid
initial conditions \( (X^1_0, \ldots, X^n_0) \) with common law \( \lambda \)
3. The mean field case (McKean-Vlasov 1966)

\[ G_n = K_n \text{ complete graph; wlog } V_n = \{1, \ldots, n\} \]

Particles \( i = 1, \ldots, n \) interact according to

\[
dX^i_t = \frac{1}{n} \sum_{k=1}^{n} b(X^i_t, X^k_t) dt + dW^i_t, \quad i = 1, \ldots, n
\]

where \( W^1, \ldots, W^n \) are independent Brownian motions, with iid initial conditions \( (X^1_0, \ldots, X^n_0) \) with common law \( \lambda \)

This can be reformulated as

\[
dX^i_t = B(X^i_t, \bar{\mu}^n_t) dt + dW^i_t, \quad \bar{\mu}^n_t = \frac{1}{n} \sum_{k=1}^{n} \delta_{X^k_t},
\]

where, for a probability measure \( m \) on \( \mathbb{R}^d \),

\[
B(x, m) := \int_{\mathbb{R}^d} b(x, y) m(dy),
\]
Theorem (McKean '67, Oelschlager '84, Sznitman '91, etc.)

\[(\overline{\mu}_t^n)_{t \in [0,T]}\] converges in probability to the unique solution of the McKean-Vlasov equation

\[dX_t = B(X_t, \mu_t)dt + dW_t, \quad \mu_t = \text{Law}(X_t),\]

with \(X_0 \sim \lambda\).

Moreover, the particles become asymptotically independent. Precisely, for fixed \(k\),

\[(X^1, \ldots, X^k) \Rightarrow \mu^\otimes k, \quad \text{as } n \to \infty.\]
The existence of a limit follows from general results on exchangeable processes:
As $n \to \infty$, $X^{(n)} \Rightarrow X^{(\infty)}$, where
\[ dX_t^{(\infty),i} = b(X_t^{(\infty),i}, \mu_t)dt + dW^i(t), \quad i = 1, 2, \ldots, \]
with
\[ \mu_t = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{X_t^{\infty},i}. \]

One then characterizes the marginal dynamics of a single particle in this infinite system.
The existence of a limit follows from general results on exchangeable processes:
As $n \to \infty$, $X^{(n)} \Rightarrow X^{(\infty)}$, where

$$dX^{(\infty),i}_t = b(X^{(\infty),i}_t, \mu_t)dt + dW^i(t), \quad i = 1, 2, \ldots,$$

with

$$\mu_t = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{X^{\infty,i}_t}.$$

One then characterizes the marginal dynamics of a single partile in this infinite system.

The McKean-Vlasov equation provides an autonomous description of this marginal dynamics.
Consider random elements \((Z^n)_{n \in \mathbb{N}}, Z\), taking values in a Polish (complete, separable, metrizable) space \(\mathcal{X}\).

**Definition of Weak Convergence**

Then \(Z^n\) is said to converge weakly to \(Z\), as \(n \to \infty\), denoted \(Z^n \Rightarrow Z\), if for each bounded, continuous function \(f : \mathcal{X} \to \mathbb{R}\),

\[
\mathbb{E}[f(Z^n)] \to \mathbb{E}[f(Z)] \quad \text{as} \quad n \to \infty.
\]

**Exercise**

Suppose \(S = \mathbb{R}\), and \(P_n, P\) lie in \(\mathcal{P}(\mathbb{R})\) and let \(F_n\) and \(F\) denote their corresponding cumulative distribution functions. Then show that \(P_n\) converges weakly to \(P\) if and only if

\[
F_n(x) \to F(x), \quad \forall x : F(x) = F(x^-).
\]
Consider random elements \((Z^n)_{n \in \mathbb{N}}, Z\), taking values in a Polish (complete, separable, metrizable) space \(\mathcal{X}\).

**Definition of Weak Convergence**

Then \(Z^n\) is said to converge weakly to \(Z\), as \(n \to \infty\), denoted \(Z^n \Rightarrow Z\), if for each bounded, continuous function \(f : \mathcal{X} \to \mathbb{R}\),

\[
\mathbb{E}[f(Z^n)] \to \mathbb{E}[f(Z)] \quad \text{as } n \to \infty.
\]

There is a metric (the Prohorov metric) on the space \(\mathcal{P}(\mathcal{X})\) of probability measures on \(\mathcal{X}\) that metrizes this notion of convergence. The resulting space is again Polish.
How does one prove convergence in general?

1. Show relative compactness of the sequence \( \{Z_n\} \).
2. Uniquely characterize any subsequential limit.
How does one prove convergence in general?

1. Show relative compactness of the sequence \( \{Z_n\} \).
2. Uniquely characterize any subsequential limit.

For weak convergence

Prohorov’s Theorem provides a criterion for relative compactness

**Definition of Tightness**

The sequence of random elements \( \{Z^n\} \) is said to be tight if for every \( \varepsilon > 0 \), there exists a compact space \( K_\varepsilon \subset X \) such that

\[
P(Z^n \notin K_\varepsilon) < \varepsilon.
\]

**Prohorov’s Theorem (1956)**

A sequence \( \{Z^n\} \) is relatively compact if and only if it is tight.
Recall

\[ dX_t^{i,n} = B(X_t^{i,n}, \bar{\mu}_t^n) \, dt + dW_t^{i,n}, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{k,n}}, \]

for \( i = 1, \ldots, n \), where \( \{W_t^{i,n}, i = 1, \ldots, n\} \) are independent \( d \)-dimensional Brownian motions and where, for a probability measure \( m \in \mathcal{P}(\mathbb{R}^d) \),

\[ B(x, m) = \int_{\mathbb{R}^d} b(x, y) \, m(dy), \]

We want to show weak convergence of \( \{\bar{\mu}_t^n, t \geq 0\} \)

Note that \( \bar{\mu}_t^n \) is a random element taking values in the (Polish) space \( X := C([0, \infty) : \mathcal{P}(\mathbb{R}^d)) \) of continuous (probability) measure-valued trajectories.
Useful in the study of Markov processes
Infinitesimal Generators

- Useful in the study of Markov processes
- The infinitesimal generator of a continuous-time stochastic process \( \{X_t, t \geq 0\} \) is a differential operator of the form:

\[
A f(x) = \lim_{t \to 0} \frac{\mathbb{E}[f(X_t) - f(x)]}{t},
\]

which acts on the set of functions \( f : \mathbb{R}^d \mapsto \mathbb{R} \) for which the above limit is well defined, called the domain of \( A \).

- \( A \) should be viewed as an “average derivative” along the flow of a stochastic process \( \{X_t, t \geq 0\} \).
Infinitesimal Generators

- Useful in the study of Markov processes
- The infinitesimal generator of a continuous-time stochastic process \( \{X_t, t \geq 0\} \) is a differential operator of the form:

\[
\mathcal{A} f(x) = \lim_{t \to 0} \frac{\mathbb{E}[f(X_t) - f(x)]}{t},
\]

which acts on the set of functions \( f: \mathbb{R}^d \to \mathbb{R} \) for which the above limit is well defined, called the domain of \( \mathcal{A} \).
- \( \mathcal{A} \) should be viewed as an “average derivative” along the flow of a stochastic process \( \{X_t, t \geq 0\} \)
- For an SDE of the form:

\[
dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x.
\]

where \( a(x) = \sigma(x)\sigma^T(x) \), \( \mathcal{A} \) takes the form

\[
\mathcal{A} f(x) = \sum_i b_i(x) \partial_if(x) + \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i \partial_j f(x)
\]
Infinitesimal Generators

- Useful in the study of Markov processes
- The infinitesimal generator of a continuous-time stochastic process \( \{X_t, t \geq 0\} \) is a differential operator of the form:

\[
Af(x) = \lim_{t \to 0} \frac{E[f(X_t) - f(x)]}{t},
\]

which acts on the set of functions \( f : \mathbb{R}^d \mapsto \mathbb{R} \) for which the above limit is well defined, called the **domain** of \( A \).

- \( A \) should be viewed as an “average derivative” along the flow of a stochastic process \( \{X_t, t \geq 0\} \)

- For an SDE of the form:

\[
dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x.
\]

where \( a(x) = \sigma(x)\sigma^T(x) \), \( A \) takes the form

\[
Af(x) = \sum_i b_i(x)\partial_i f(x) + \frac{1}{2} \sum_{ij} a_{ij}(x)\partial_i \partial_j f(x)
\]

- The **domain** of \( A \) includes \( C_b(\mathbb{R}^d) \), the space of bounded continuous real-valued functions on \( \mathbb{R}^d \).
Step 1: Show tightness of \( \{ \bar{\mu}^n \} \) \( n \in \mathbb{N} \)

For this, use the “infinitesimal generator” of the original process

\[
dX_t^{i,n} = B(X_t^{i,n}, \bar{\mu}_t^n) dt + \sigma(X_t^{i,n}, \bar{\mu}_t^n) dW_t^{i,n}, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^k,n},
\]

Set \( a(x, \mu) := \sigma(x, \mu) \sigma^T(x, \mu) \), where \( \mu = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \), and fix

\[
g \in C^2_b(\mathbb{R}), \quad h \in C^2_b(\mathbb{R}^d), \quad \langle \mu, h \rangle = \int_{\mathbb{R}} h(x) \mu(dx)
\]

For \( f(\mu) = g(\langle \mu, h \rangle) \), define

\[
\mathcal{A}_h^n(\mu) = g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}
\]

\[
+ \frac{1}{2} g''(\langle \mu, h \rangle) \frac{1}{n} \langle \mu, a_{ij}(\cdot, \mu) \partial_i h \partial_j h \rangle.
\]

Use the generator, martingale estimates and stochastic calculus (Itô’s formula) to show that \( \{ \bar{\mu}^n \} \) \( n \in \mathbb{N} \) is tight;
Recall that the evolution of $\mu^n$ is characterized by

$$A^n_h(\mu) = g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}$$

$$+ \frac{1}{2} g''(\langle \mu, h \rangle) \frac{1}{n} \langle \mu, a_{ij}(\cdot, \mu) \partial_i h \partial_j h \rangle.$$
Recall that the evolution of $\bar{\mu}^n$ is characterized by

$$A^n_h(\mu) = g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}$$

$$+ \frac{1}{2} g''(\langle \mu, h \rangle) \frac{1}{n} \langle \mu, a_{ij}(\cdot, \mu) \partial_i h \partial_j h \rangle.$$ 

**Step 2:** Show that any subsequential limit $\bar{\mu}$ of $\{\mu^n\}_{n \in \mathbb{N}}$ satisfies a certain martingale problem: or more precisely, satisfies

$$A^\infty_h(\bar{\mu}) = 0, \quad \forall h \in C_b(\mathbb{R}^d),$$

where

$$A^\infty_h(\mu) := g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}$$

Note that $A^\infty(\mu)$ is a nonlinear differential operator.
Outline of Proof of Mean-Field Limit (contd.)

Recall

- **Step 1:** Show \( \{\bar{\mu}^n\} \) is tight, and
- **Step 2:** Every subsequential limit \( \bar{\mu} \) of \( \{\bar{\mu}^n\} \) satisfies

\[
\mathcal{A}_{h}^{\infty}(\bar{\mu}) = 0, \quad \forall h \in \mathcal{C}_b(\mathbb{R}^d),
\]

where

\[
\mathcal{A}_{h}^{\infty}(\mu) := g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}
\]
Recall

- **Step 1:** Show $\{\bar{\mu}^n\}$ is tight, and

- **Step 2:** Every subsequential limit $\bar{\mu}$ of $\{\bar{\mu}^n\}$ satisfies

  $$\mathcal{A}_h^\infty(\bar{\mu}) = 0, \quad \forall h \in C_b(\mathbb{R}^d),$$

  (1)

  where

  $$\mathcal{A}_h^\infty(\mu) := g'(\langle \mu, h \rangle) \left\{ \langle \mu, b(\cdot, \mu) \cdot \nabla h \rangle + \frac{1}{2} \langle \mu, a_{ij}(\cdot, \mu) \partial_{ij} h \rangle \right\}$$

- **Step 3:** Show uniqueness of solutions to the (weak) PDE (29) to conclude that $\{\bar{\mu}^n\}$ converges to the unique (weak) solution $\bar{\mu}$ of the PDE (29).
Outline of Proof of Mean-Field Limit (contd.)

\[ dX_t^{i,n} = B(X_t^{i,n}, \bar{\mu}_t^n)dt + \sigma(X_t^{i,n}, \bar{\mu}_t^n)dW_t^{i,n}, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_{t}^{k,n}}, \]

**Steps 1–3:** We have shown that \( \bar{\mu}^n \to \bar{\mu} \), the unique (weak) solution to

\[ \mathcal{A}_h^\infty(\bar{\mu}) = 0, \quad \forall h \in C_b(\mathbb{R}^d), \]

\[ K. Ramanan \quad Scaling \ Limits \]
\[ dX_t^{i,n} = B(X_t^{i,n}, \bar{\mu}_t^n)dt + \sigma(X_t^{i,n}, \bar{\mu}_t^n)dW_t^{i,n}, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{k,n}}. \]

**Steps 1–3:** We have shown that \( \bar{\mu}_t^n \to \bar{\mu} \), the unique (weak) solution to
\[ A_\infty^n(\bar{\mu}) = 0, \quad \forall h \in C_b(\mathbb{R}^d), \]

**Step 4:** We combine this with the above SDE for \( X^{i,n} \) and the continuity of the map \( m \mapsto B(x, m) \) to conclude that as \( n \to \infty \), for any \( i \), \( X^{i,n} \) converges weakly to \( X \), which is the unique solution to the (nonlinear or McKean-Vlasov) SDE:
\[ dX_t = B(X_t, \bar{\mu}_t)dt + \sigma(X_t, \bar{\mu}_t)dW_t, \]
for some Brownian motion \( (W_t)_{t \geq 0} \).

**Step 5:** Noting that the PDE for \( \mu \) is the forward Kolmogorov equation for \( X \), we conclude that \( \text{Law}(X_t) = \mu_t \).
Theorem (McKean '67, Oelschlager '84, Sznitman '91, etc.)

\((\bar{\mu}^n(t))_{t\in[0,T]}\) converges in probability to the unique solution
\((\mu(t))_{t\in[0,T]}\) of the McKean-Vlasov equation

\[dX_t = B(X_t, \mu_t)dt + dW_t, \quad \mu_t = \text{Law}(X_t),\]

\(X_0 \sim \lambda\). Here, \(\mu_t\) satisfies a nonlinear PDE that is the forward Kolmogorov equation for this inhomogeneous Markov process.
An alternative approach: asymptotic independence

Recall

\[ dX_t^{n,i} = B(X_t^{n,i}, \bar{\mu}_t^n)dt + dW_t^i, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{n,k}}, \]

for \( i = 1, \ldots, n \), where, for a probability measure \( m \) on \( \mathbb{R}^d \),

\[ B(x, m) = \int_{\mathbb{R}^d} b(x, y) m(dy), \]
An alternative approach: asymptotic independence

Recall

\[ dX_t^{n,i} = B(X_t^{n,i}, \bar{\mu}_t^n)dt + dW_t^i, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{n,k}}, \]

for \( i = 1, \ldots, n \), where, for a probability measure \( m \) on \( \mathbb{R}^d \),

\[ B(x, m) = \int_{\mathbb{R}^d} b(x, y) m(dy), \]

Note: If \( X^{n,i}, i = 1, \ldots, n, n \in \mathbb{N} \), were independent,
An alternative approach: asymptotic independence

Recall

\[ dX^n_{t,i} = B(X^n_{t,i}, \bar{\mu}^n_t)dt + dW^i_t, \quad \bar{\mu}^n_t = \frac{1}{n} \sum_{k=1}^{n} \delta_{X^n_{t,k}}, \]

for \( i = 1, \ldots, n \), where, for a probability measure \( m \) on \( \mathbb{R}^d \),

\[ B(x, m) = \int_{\mathbb{R}^d} b(x, y) m(dy), \]

Note: If \( X^n_{i}, i = 1, \ldots, n, n \in \mathbb{N} \), were independent, the SLLN would tell us that \( \bar{\mu}^n_t \to \text{Law}(X^i_t) \)

What we have here is a weak sort of dependence – can show asymptotic independence holds, to conclude
An alternative approach: asymptotic independence

Recall

\[ dX_t^{n,i} = B(X_t^{n,i}, \bar{\mu}_t^n)dt + dW_t^i, \quad \bar{\mu}_t^n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_t^{n,k}}, \]

for \( i = 1, \ldots, n \), where, for a probability measure \( m \) on \( \mathbb{R}^d \),

\[ B(x, m) = \int_{\mathbb{R}^d} b(x, y) m(dy), \]

**Note:** If \( X^{n,i}, i = 1, \ldots, n, n \in \mathbb{N} \), were independent, the SLLN would tell us that \( \bar{\mu}_t^n \to \text{Law}(X_t^i) \)

What we have here is a weak sort of dependence – can show asymptotic independence holds, to conclude

**Theorem**

\( (\bar{\mu}_t^n(t))_{t \in [0, T]} \) converges in probability to the unique solution
\( (\mu_t)_{t \in [0, T]} \) of the McKean-Vlasov equation

\[ dX_t = B(X_t, \mu_t)dt + dW_t, \quad \mu_t = \text{Law}(X_t). \]
Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, how can we describe the limiting behavior of a “typical” particle $X_t^\nu$?
Key questions

Given a sequence of graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, how can we describe the limiting behavior of a “typical” particle $X_t^v$?

Theorem (Delattre-Giacomin-Luçon ’16; Bhamidi-Budhiraja-Wu ’19)

Under suitable conditions on the coefficients, suppose $G_n = G(n, p_n)$ is Erdős-Rényi, with $np_n \to \infty$. Then everything behaves like in the mean field case.

See also Delarue ’17, Coppini, Dietert and Giacomin ’18, Reis and Oliveira ’18

Observation: $np_n \approx$ average degree, so $np_n \to \infty$ means the graphs are suitably dense.
The Main Focus of this Lecture Series

the sparse graph regime
The Main Focus of this Lecture Series

the sparse graph regime

In this regime, there was not even an existing conjecture as to:

(i) whether it is possible to characterize the limiting dynamics of a typical particle
(ii) what form this characterization would take
6. Beyond Mean-Field Limits ...

\[ X_t^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_{t+1}^v \right), \quad v \in V, \]

\[ dX_t^{G_n,v} = \frac{1}{d_v} \sum_{u \sim v} b(X_t^{G_n,v}, X_t^{G_n,u})dt + dW^v(t) \]

**Key questions**

Given a sequence of graphs \( G_n = (V_n, E_n) \) with \(|V_n| \to \infty\), how can we describe the limiting behavior of...

- the state of a “typical” or fixed particle \( X_t^{G_n,v} \)?
- the empirical measure of particles \( \mu_t^{G_n} := \frac{1}{|V_n|} \sum_{v \in V_n} \delta_{X_t^{G_n,v}} \)?

**Our focus:** The sparse regime, where degrees do not diverge. How does the \( n \to \infty \) limit reflect the graph structure?

**Example:** Erdős-Rényi \( G(n, p_n) \) with \( np_n \to p \in (0, \infty) \).

**Open Question:** in Delattre-Giacomin-Luçon
6. Beyond Mean-Field Limits ...

\[ X_{t+1}^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_t^v \right), \ v \in V, \]

\[ dX_{t}^{G_n,v} = \frac{1}{d_v} \sum_{u \sim v} b(X_t^{G_n,v}, X_t^{G_n,u}) dt + dW^v(t), \quad v \in V \]
6. Beyond Mean-Field Limits ...

\[ X_{t+1}^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_t^{v+1} \right), \; v \in V, \]

\[ dX_t^G_{n,v} = \frac{1}{d_v} \sum_{u \sim v} b(X_t^G_{n,v}, X_t^G_{n,u}) \, dt + dW^v(t), \; \quad v \in V \]

Specific Questions:
(1) Does the whole system admit a scaling limit?
(2) Is there an autonomous description of the limiting dynamics of a typical particle?
(3) Does the (global) empirical measure

\[ \bar{\mu}^G_n = \frac{1}{|V_n|} \sum_{v \in V_n} \delta_{X^G_{n,v}(\cdot)} \]

have a scaling limit?
Sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$, 

\[ X^v_{t+1} = F \left( X^v_t, (X^u_t)_{u \sim v}, \xi^v_{t+1} \right), \quad v \in V, \]

\[ dX^G_{n,v} = \frac{1}{d_v} \sum_{u \sim v} b(X^G_{n,v}, X^G_{n,u}) dt + dW^v(t), \quad v \in V \]
Sequence of sparse graphs $G_n = (V_n, E_n)$ with $|V_n| \to \infty$,

$$X_{t+1}^v = F \left( X_t^v, (X_t^u)_{u \sim v}, \xi_{t+1}^v \right), \quad v \in V,$$

$$dX_{t}^{G_n,v} = \frac{1}{d_v} \sum_{u \sim v} b(v, X_{t}^{G_n,v}, X_{t}^{G_n,u}) dt + dW^v(t), \quad v \in V.$$

(Q1) Does the whole system admit a scaling limit?

(A1) Yes, wrt a generalized notion of local weak convergence

In fact, for this, we can allow more general heterogeneous dynamics: e.g.,

$$X_{t+1}^v = F_v \left( t, X_t^v, (X_t^u)_{u \sim v}, \xi_{t+1}^v \right), \quad v \in V,$$

or

$$dX_{t}^{G_n,v} = \frac{1}{d_v} \sum_{u \sim v} b_v(t, X_{t}^{G_n,v}, X_{t}^{G_n,u}) dt + dW^v(t), \quad v \in V.$$