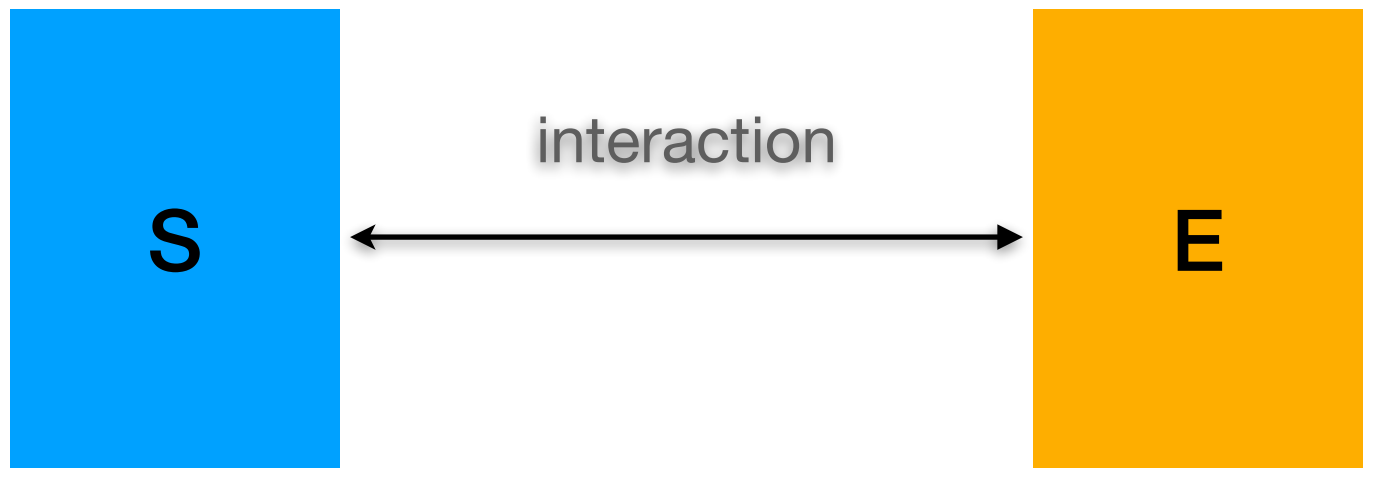
# Controlled Dynamical Systems in the Space of Probability Measures

## IMSI Workshop on New Directions in Reinforcement Learning and Control

Maxim Raginsky (UIUC)

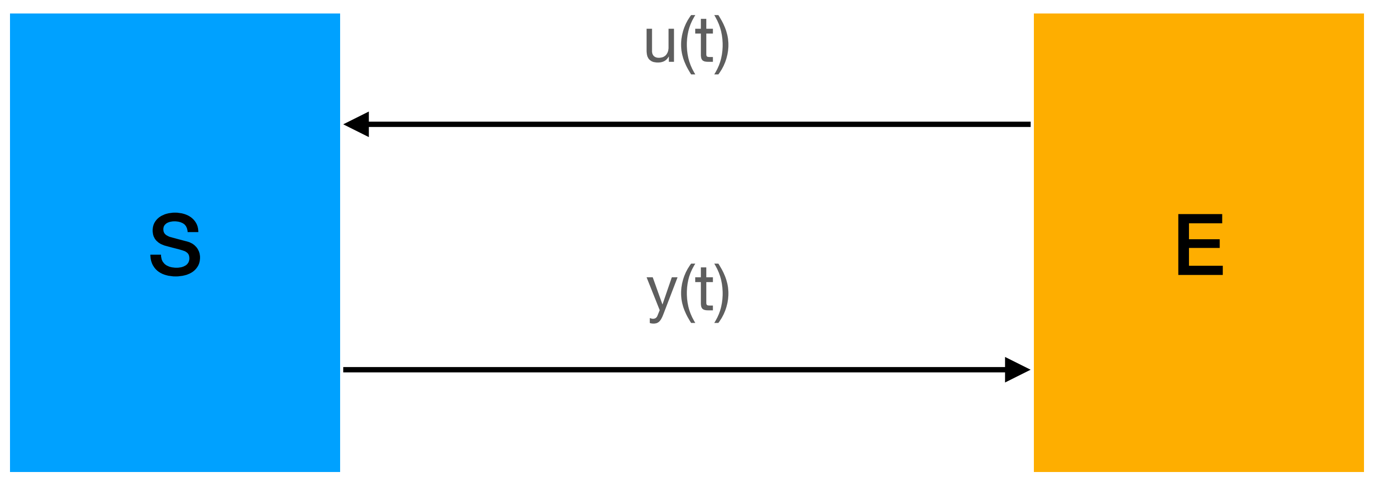
# Some Generalities on Systems

- Engineering is about synthesis of systems with desired function, behavior, or properties.
- Unlike (most of) physics, engineering deals with systems that are not isolated from their environment, but are interacting with it.
    - Separation into system and environment is context- and application-dependent.
- The behavior of the system can be observed (perhaps inaccurately), modeled, and changed by a feedback control mechanism.



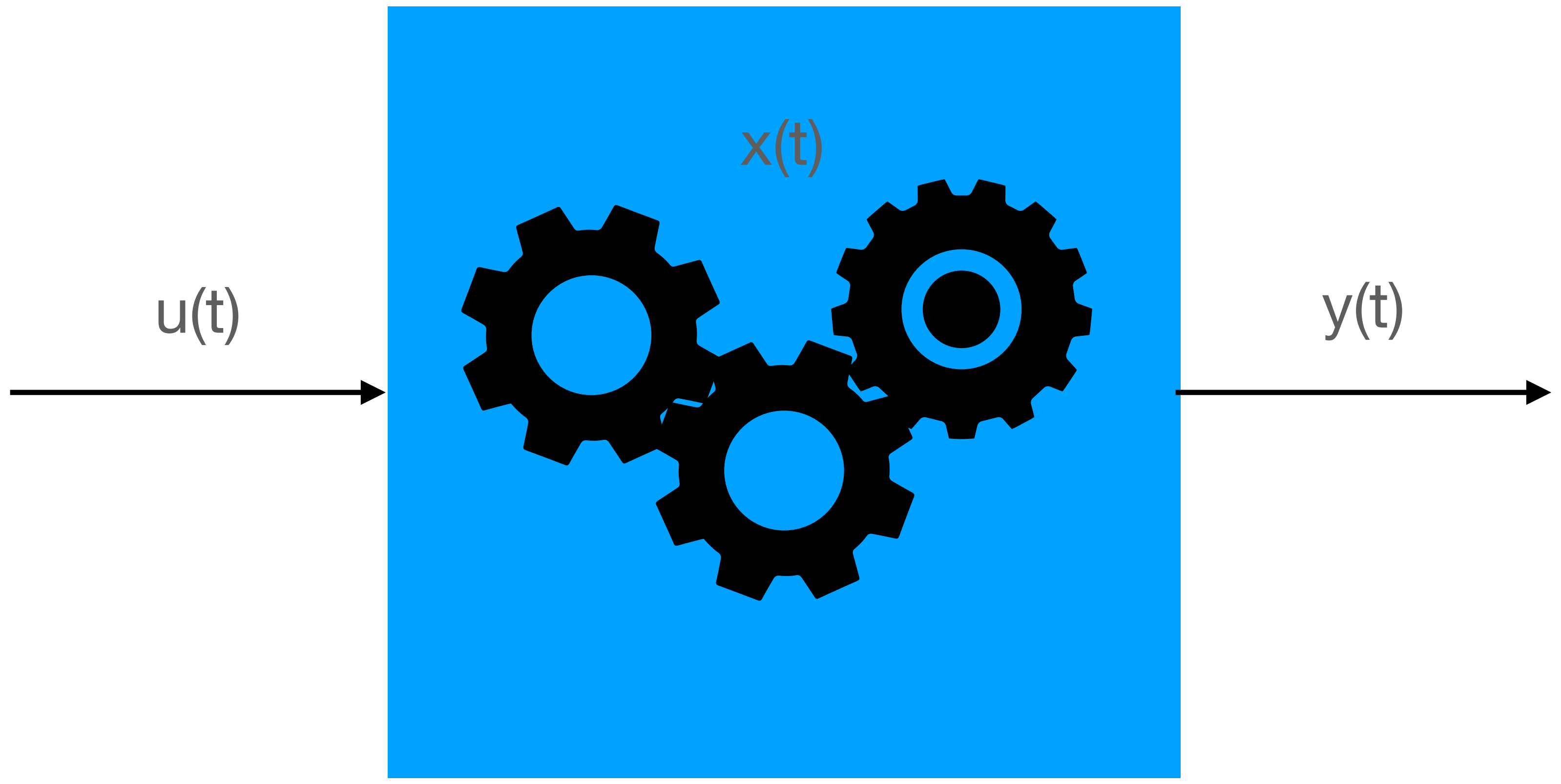
# The Notion of State

- Externalist perspective: inputs are imposed on the system from the "outside," outputs are generated in response.




- Internalist perspective: the state as system memory

$$\dot{x} = f(x, u), \quad y = g(x, u)$$




- Autonomous systems: $\dot{x} = f(x)$ (the state of the system at time t completely determines the trajectory for all times s > t).

- Realization problem: reconciling externalist and internaliist perspectives.

# Controlled Dynamical Systems

**state, control, trajectory**

- Dynamical systems have attributes that can be observed externally and that evolve in time
- Controlled dynamical system (CDS): evolution can be shaped by imposing inputs (controls)
- State = system memory; knowledge of the state + future inputs determines future trajectory
- Standard questions about a CDS:
  - well-posedness — is it possible to associate a unique trajectory with an initial state and a control law?
  - controllability — is it possible to steer the system from a given initial state to a given target state?
  - optimal control — among all controlled trajectories satisfying a given constraint, select one that's (nearly) the best

# Controlled Dynamics on $\mathscr{P}(\mathbb{R}^n)$

- State space: space of probability measures on $\mathbb{R}^n$
- Function of control inputs: influence the evolution of probability-valued state
- Model template:
  $\dot{\mu} = F[\mu, u]$ — $\mu \in \mathscr{P}(\mathbb{R}^n)$: state, $u \in U$: control input
- Applications:
  - generation/sampling
  - belief updating (e.g., in partially observed stochastic control problems)
  - beautiful mathematics

# Controlled Ordinary Differential Equations

- System model:

  $\dot{x} = f(x, u), x(0) = x_0, t \in [0,T] \qquad x \in \mathbb{R}^n, u \in U$

- Taxonomy of control laws:
  - open loop — $\gamma : [0,T] \to U,\ u(t) = \gamma(t)$
  - state feedback: $\varphi : [0,T] \times \mathbb{R}^n \to U,\ u(t) = \varphi(t, x(t))$
- Optimality:

$$\text{minimize} \quad J(x_0; u(\cdot)) = \int_0^T q(t, x(t), u(t))dt + r(x(T))$$

# Controlled Ordinary Differential Equations

- System model:
  $\dot{x} = f(x, u), x(0) = x_0, t \in [0,T] \qquad x \in \mathbb{R}^n, u \in U \subseteq \mathbb{R}^m$
- Random initial state: $x(0) \sim \mu_0$
- Evolution of probability measures on $\mathbb{R}^n$: controlled continuity equation

$$\mu_t := \mathscr{L}(x(t))$$
$$\partial_t \mu_t = -\nabla \cdot (f(\cdot, u(\cdot))\mu_t), \qquad t \in [0,T]$$

- Interpretation: for any $h \in C^1(\mathbb{R}^n, \mathbb{R})$,

$$\langle \mu_t, h \rangle = \langle \mu_0, h \rangle + \int_0^t \langle \mu_s, f(\cdot, u(s))^T \nabla h \rangle ds \qquad \longleftrightarrow \qquad \dot{\mu} = F[\mu, u]$$

# Controlled Stochastic Differential Equations

- System model:
  $dX_t = b(X_t, u_t)dt + \sigma(X_t, u_t)dW_t, \quad X_0 \sim \mu_0, t \in [0,T]$
- Evolution of probability measures on $\mathbb{R}^n$: controlled forward Kolmogorov equation

$$\partial_t \mu_t = -\nabla \cdot (b(\cdot, u_t)\mu_t) + \frac{1}{2}\mathrm{tr}\left[\nabla^2(\sigma(\cdot, u_t)\sigma(\cdot, u_t)^T \mu_t)\right], \qquad t \in [0,T]$$

- Interpretation: for any $h \in C^2(\mathbb{R}^n, \mathbb{R})$,

$$\langle \mu_t, h \rangle = \langle \mu_0, h \rangle + \int_0^t \langle \mu_s, L_s^{u(\cdot)} h \rangle ds \qquad \longleftrightarrow \qquad \dot{\mu} = F[\mu, u]$$

$$L_s^{u(\cdot)} h(x) := b(x, u_s)^T \nabla h(x) + \frac{1}{2}\mathrm{tr}\left[\sigma(x, u_s)\sigma(x, u_s)^T \nabla^2 h(x)\right]$$

# Dynamical Systems on $\mathscr{P}(\mathbb{R}^n)$ — No Control Yet

**D.W. Stroock, *Markov Processes from K. Itô's Perspective***

- What do we mean by dynamics on $\mathscr{P}(\mathbb{R}^n)$?
- *Differentiable curve* in $\mathscr{P}(\mathbb{R}^n)$: $t \mapsto \mu_t, t \geq 0$
  - there exists, for each $t$, a continuous linear functional $\dot{\mu}_t$ on $C^\infty(\mathbb{R}^n, \mathbb{R})$ such that the map $(t, h) \mapsto \dot{\mu}_t h$ is continuous and

$$\langle \mu_t, h \rangle = \langle \mu_0, h \rangle + \int_0^t \dot{\mu}_s h \, ds, \qquad \text{for all } t \geq 0$$

- In analogy to dynamical systems on $\mathbb{R}^n$, we would like to have a notion of a *vector field* on $\mathscr{P}(\mathbb{R}^n)$, so that we can write something like $\dot{\mu} = F[\mu]$

# A Building Block: Lévy Systems

**drift, diffusion, jumps**

- A *Lévy system* is a triple $(b, a, M)$, where $b \in \mathbb{R}^n$ is a vector, $a \in \mathbb{R}^{n \times n}$ is a symmetric and positive semidefinite matrix, and $M$ is a Lévy measure, i.e., positive Borel measure on $\mathbb{R}^n$ that satisfies $M(\{0\}) = 0$ and

$$\int_{\mathbb{R}^n} \frac{|y|^2}{1 + |y|^2} M(dy) < \infty$$

- Lévy systems are building blocks of Lévy processes — stochastic processes with stationary and independent increments that exhibit three basic forms of stochastic behavior: drift, diffusion, and jumps
- Examples: Brownian motion (possibly with drift), Poisson and compound Poisson processes

# Vector Fields on $\mathscr{P}(\mathbb{R}^n)$

## the role of Lévy-Khintchine operators

- Stroock's definition:
  A *vector field $F[\,\cdot\,]$ on $\mathscr{P}(\mathbb{R}^n)$* is a mapping $(x, \mu) \mapsto (b(x, \mu), a(x, \mu), M((x, \mu), \cdot\,))$ that associates a Lévy system to each $x \in \mathbb{R}^n$ and $\mu \in \mathscr{P}(\mathbb{R}^n)$, such that, for every $h \in C_b^2(\mathbb{R}^n, \mathbb{R})$,

  $$F[\mu]h = \langle L_\mu h, \mu \rangle$$

  where $L_\mu h(x) := b(x, \mu)^T \nabla h(x) + \frac{1}{2}\mathrm{tr}[a(x, \mu)\, \nabla^2 h(x)]$

  $$+ \int_{\mathbb{R}^n} \left( h(y + x) - h(x) - \frac{\langle \nabla h(x), y \rangle}{1 + |y|^2} \right) M((x, \mu), dy)$$

- Note: We will keep things simple and will only deal with vector fields where $M \equiv 0$ (no jumps).

# Dynamical Systems in $\mathscr{P}(\mathbb{R}^n)$

**well-posedness and the question of realization**

- Let a dynamical system on $\mathscr{P}(\mathbb{R}^n)$ be given: $\dot{\mu}_t = F[\mu_t], \mu_0 = \mu$
- *Well-posedness:* does this measure-valued ODE have a unique solution $\mu_t$?

$$\langle \mu_t, h \rangle = \langle \mu, h \rangle + \int_0^t \langle \mu_s, L_{\mu_s} h \rangle ds, \qquad t \geq 0$$

where $L_\mu$ is the Lévy-Khintchine operator associated with $F[\mu]$

- *Realization:* can we construct a stochastic process $X = (X_t)_{t \geq 0}$, such that $\mu_t = \mathscr{L}(X_t)$ and the sample paths of $X$ inherit the continuity properties of $t \mapsto \mu_t$?

# Well-Posedness on $\mathscr{P}(\mathbb{R}^n)$

- Let $F[\,\cdot\,]$ be a vector field on $\mathscr{P}(\mathbb{R}^n)$ with $M \equiv 0$, such that, for all $x, x'$ and $\mu, \nu$

  $$|\,b(x,\mu) - b(x',\nu)\,| + \|\sigma(x,\mu) - \sigma(x',\nu)\| \leq K(\,|\,x - x'\,| + W_2(\mu,\nu)),$$

  where $\sigma(\,\cdot\,,\cdot\,) = \sqrt{a(\,\cdot\,,\cdot\,)}$ and where $W_2(\,\cdot\,,\cdot\,)$ is the 2-Wasserstein distance on $\mathscr{P}(\mathbb{R}^n)$.
- Then, for any $\mu \in \mathscr{P}(\mathbb{R}^n)$ with a finite 2nd moment, the ODE $\dot{\mu}_t = F[\mu_t]$, $t \in [0,T]$, with initial condition $\mu_0 = \mu$ has a unique solution.
- Proof idea:
  - Existence — Euler approximation + equicontinuity of approximations in 2-Wasserstein.
  - Uniqueness — a fixed-point argument, goes through process realization (up next).

# Process Realization on $\mathbb{R}^n$

- Let $F[\,\cdot\,]$ be a vector field on $\mathscr{P}(\mathbb{R}^n)$ satisfying our regularity conditions.
- Consider the Itô SDE of McKean-Vlasov type:

$$dX_t = b(X_t, \mathscr{L}(X_t))dt + \sigma(X_t, \mathscr{L}(X_t))dW_t, \qquad X_0 \sim \mu; t \in [0,T]$$

  Then, for all $t \in [0,T]$, $X_t \sim \mu_t$ where $\dot{\mu}_t = F[\mu_t]$.
- Proof idea:
  - For any "exogenous" curve $t \mapsto \nu_t \in \mathscr{P}(\mathbb{R}^n)$ with $\nu_0 = \mu$, the SDE with time-dependent drift $b(x, \nu_t)$ and diffusion matrix $\sigma(x, \nu_t)$ has a unique strong solution $(X_t^\nu)_{t\in[0,T]}$.
  - By the contraction mapping principle on Wasserstein space of paths, the map $(\nu_t)_{t\in[0,T]} \mapsto (\mathscr{L}(X_t^\nu))_{t\in[0,T]}$ has a unique fixed point, which must coincide with $(\mu_t)_{t\in[0,T]}$.

# Time-Varying Dynamics

- Handling time-dependent vector fields $F[t, \cdot\,]$ is straightforward: $(t, x, \mu) \mapsto (b(t, x, \mu), a(t, x, \mu))$

- Assume Lipschitz continuity uniformly in time: $|\,b(t, x, \mu) - b(t, x', \nu)\,| + \|\sigma(t, x, \mu) - \sigma(t, x', \nu)\| \leq K(\,|\,x - x'\,| + W_2(\mu, \nu)), \qquad \text{for all } t \in [0,T]$

- Then the time-varying system $\dot{\mu}_t = F[t, \mu_t]$ has a unique solution for any initial condition $\mu \in \mathscr{P}(\mathbb{R}^n)$ with finite second moments, and admits a process realization $(X_t)_{t\in[0,T]}$ governed by the McKean-Vlasov SDE
$dX_t = b(t, X_t, \mathscr{L}(X_t))dt + \sigma(t, X_t, \mathscr{L}(X_t))dW_t, \quad X_0 \sim \mu; t \in [0,T]$

# Adding Controls

- A controlled dynamical system on $\mathscr{P}(\mathbb{R}^n)$ is a collection of vector fields $F[\mu, u]$ indexed by the elements $u$ of a control set $U$.
- Just like in the case of $\mathbb{R}^n$, we can distinguish different types of control laws:
  - open loop — $u_t$ depends only on $t$
  - (time-dependent) state feedback — given $(b(x, u, \mu), a(x, u, \mu))$, let $u_t = \phi(t, x)$
  - (time-dependent) measure feedback — now let $u_t$ depend on $t, x, \mu$
- Need to impose appropriate conditions on the control law to guarantee well-posedness.

# Affine Vector Fields

- Following Stroock, a vector field $F$ on $\mathscr{P}(\mathbb{R}^n)$ is *affine* if
$F[\theta\mu + (1-\theta)\nu] = \theta F[\mu] + (1-\theta)F[\nu], \qquad \mu, \nu \in \mathscr{P}(\mathbb{R}^n), \theta \in [0,1]$
- Equivalently, $F$ is affine iff its Lévy system does not depend on $\mu$. These correspond to diffusion processes of the usual kind (no McKean-Vlasov nonlinear behavior):
$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$
- Controlled affine vector fields: $F[t, \cdot, u]$ is affine for each $u \in U$:
$dX_t = b(t, X_t, u_t)dt + \sigma(t, X_t, u_t)dW_t$
- Depending on the type of control law (open loop, state feedback, measure feedback), we will end up with different types of process realizations.

# Example 1: Optimal Transportation

- Take $b(x, u) = u, a(x, u) = 0$.
- Optimal control problem in the space of densities, with controls in state feedback form:

$$\text{minimize } J(u) := \int_0^1 \langle \mu_t, |v(t, \cdot)|^2 \rangle dt$$

$$\text{subject to } \begin{array}{l} \partial_t \mu_t = -\nabla \cdot (v(t, \cdot)\mu_t), \qquad t \in [0,1] \\ \mu_0(dx) = \rho_0(x)dx, \mu_1(dx) = \rho_1(x)dx \end{array}$$

- The minimum cost is given by

$$\min_{u(\cdot)} J(u) = W_2^2(\mu_0, \mu_1)$$

(Benamou-Brenier)

# Example 2: Mean-Field Control Problem

- Take $b(x, u) = u, a(x, u) = I_n$.
- Optimal control problem in the space of densities:

$$\text{minimize } J(u) := \frac{1}{2}\int_0^1 \langle \mu_t, |v(t,\cdot)|^2 \rangle dt + r(\mu_1), \qquad \mu_0 = \delta_0$$

where $r(\mu) := 2\langle f, \mu \rangle + \langle K, \mu \otimes \mu \rangle$, $K : \mathbb{R}^n \times \mathbb{R}^n \to [0,\infty)$ symmetric.

- Controlled dynamics on $\mathscr{P}(\mathbb{R}^n)$:

$$\partial_t \mu_t = -\nabla \cdot (v(t,\cdot)\mu_t) + \frac{1}{2}\Delta \mu_t, \qquad \mu_0 = \delta_0$$

- Process realization: $dX_t = v(t, X_t)dt + dW_t, X_0 = 0,\ t \in [0,1]$

# Example 2: Mean-Field Control Problem

- ***Theorem*** *(Tzen-Raginsky)* Suppose there exists a unique solution $\mu_\star$ of the Boltzmann fixed-point problem

$$\frac{d\mu_\star}{d\gamma_n}(x) \propto \exp\left(-S(x,\mu_\star)\right), \qquad S(x,\mu) := f(x) + \int_{\mathbb{R}^n} K(x,x')\mu(dx'),$$

where $\gamma_n$ is the standard Gaussian measure on $\mathbb{R}^n$. The following state feedback control law is optimal:

$$v(t,x) = -\nabla \log \mathbf{E}[\exp(-S(W_1,\mu_\star)) \,|\, W_t = x],$$

where $W = (W_t)_{t\in[0,1]}$ is the standard Brownian motion on $\mathbb{R}^n$. The resulting flow of measures $(\mu_t)_{t\in[0,1]}$ attains $\mu_1 = \mu_\star$ and solves the forward Kolmogorov equation

$$\partial_t \mu_t = -\nabla\cdot(v(t,\cdot)\mu_t) + \frac{1}{2}\Delta\mu_t, \qquad \mu_0 = \delta_0.$$

# Example 3: Transformer Dynamics

- Let the control space be $U := (\mathbb{R}^{n\times n})^2 \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^n)$.
- Take $a(x,u) = 0$ and, for $u = (A, B, w, \nu)$,

  $$b(x, (A, B, w, \nu)) := \mathbf{P}_x^{\perp}\left(\mathscr{A}(x, A, B, \nu) + f(x, w)\right)$$

  where $\mathscr{A}(x, A, B, \nu) := \dfrac{\int e^{x^T A y} B y \nu(dy)}{\int e^{x^T A z} \nu(dz)}$ is the attention map and $f(\,\cdot\,, w)$ is a NN with weights $w$.
- We get transformer dynamics on the sphere $S^{n-1}$ by letting $\nu_t = \mu_t$ in

  $$\dot{\mu}_t = F[\mu_t, (A_t, B_t, w_t, \nu_t)]$$

  This can then be analyzed using methods developed by Geshkovski, Polyanskiy, Rigollet, etc.

# Thank You!