Outline

1) Introduction to finite-state MFGs.
   i) N-agent games with interactions.
   ii) Limiting behavior and MFGs.
   iii) MFG Equilibria.
   iv) HJB equation + Kolmogorov Fokker–Planck equation.
2) The Master Equation.
   i) Heuristic Derivation.
   ii) Probabilistic Meaning + Propagation of Chaos.
   iii) Regularity and Its Consequences.
3) Learning–Based Methods.
   i) DGME.
   ii) Operator Learning on HJB–KFP system.

References

- [BC '18] : arXiv: 1707.02648.
- [CP '19] : arXiv: 17.07.01819.
- [Carmona + Delarue '18]: Prob. Theory of MFGs w/ Applications I.

Introduction to Finite State MFGs

Finite state MFGs, in our setting, are finite-time horizon, continuous-time games, expressing the limit of such an n-player game w/ weak interactions as $n \to \infty$. When $n < \infty$, we take the following setup:

- State space is $[d] = \{1, \dots, d\}$.
- Player i's state at time t is X_t^i a process on $(\Omega, \mathcal{F}, \mathbb{P})$
- Player i's action at time t is α_t^i, and

$$\alpha_t = (\alpha_t^1, \dots, \alpha_t^n)$$

is an adapted process to the filtration generated by X_t.

Ex: Cybersecurity; states are $\{DU, DI, SU, SI\}$, and players may choose to be protected / unprotected at time t, incurring a cost for protection + infection.

More formally, take a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \geq 0}$. Players' controls take the form of an adapted process $\alpha_t = (\alpha_t^1, \dots, \alpha_t^n)$. The players' states

$X_t = (X_t^1, \dots, X_t^n)$ are characterized by a continuous-time Markov process, and $\alpha_t^i = \phi^i(t, X_t)$ for some deterministic, measurable ϕ^i on $[0, T] \times [d]^n$. Transitions may also depend on

$$\mu_t^{n,i} = \frac{1}{n-1} \sum_{j \neq i} \delta_{X_t^j} \in \underbrace{\mathcal{P}([d])}_{(d-1)\text{-dim. probability simplex}},$$

the empirical dist of all other agents.

def: A matrix $Q \in \mathbb{R}^{d \times d}$ is a Q-matrix if

generator of cont. time, finite state Markov process

i) $q(x, y) \geq 0, \quad x, y \in [d], \ x \neq y$.

ii) $q(x, x) = -\sum_{x \neq x'} q(x, x'), \quad x \in [d]$.

Now, players transition states according to rates

$$(\lambda_t(x, x', \mu, \alpha))_{x, x' \in [d]},$$

which are continuous (in the weak topology), Q-matrices, and satisfy

$$|\lambda_t(x, x', \mu, \alpha))| \leq c_\lambda(1 + |\alpha|).$$

Then, players transition according to:

$$\mathbb{P}[X_{t+h} = \underline{x} \mid X_t = \underline{y}] =$$

$$\begin{cases} 1 - \sum_{j=1}^{n} \lambda_t(x^j, y^j, \mu_t^{n,j}, \overset{\alpha_t^j}{\phi^j(t, \underline{x})})h + o(h), & \underline{x} = \underline{y}, \\ \lambda_t(x^i, y^i, \mu_t^{n,i}, \overset{\alpha_t^j}{\phi^i(t, \underline{x})})h + o(h), & x_i \neq y_i, \\ & x_j = y_j,\ j \neq i \end{cases}$$

As $h \to 0^+$, only one player can transition at a time! This is important because of dependence on $\mu_t^{n,i}$, the emp. dist.

Player i seeks to minimize a cost

$$J^i(\alpha^1, \ldots, \alpha^n) = \mathbb{E}\left[\int_0^T f(t, X_t^i, \mu_t^{n,i}, \alpha_t^i)\,dt + g(X_T^i, \mu_T^{n,i})\right],$$

where α^i are admissible controls.

From the symmetric n player formulation we pass to the mean-field limit in hopes of obtaining a simpler model. Now, for a deterministic flow of measures $(\mu_t) \geq 0$, consider the cost

$$J^\mu(\alpha) = \mathbb{E}\left[\int_0^T f(t, X_t, \mu_t, \alpha_t)\,dt + g(X_T, \mu_T)\right],$$

where $(X_t)_{t\geq 0}$ is a nonhomogeneous Markov chain with transition probabilities determined by

$$q_t^{\mu,\alpha}(x,y) = \lambda_t(x, y, \mu_t, \alpha_t).$$

It is then natural to seek the flow $\mu = (\mu_t)_{t\geq 0}$ s.t. the flow of the law of $(X_t)_{t\geq 0}$ solves the preceding control problem, i.e., $\mathcal{L}(X_t) = \mu_t$. This, informally, is the type of equilibrium we want to characterize via a PDE. First, if L_t is the infinitesimal generator of X_t, then

$$\begin{aligned}
[L_t p](x) &= q_t(x,\cdot)p(\cdot) \\
&= \sum_{y\in[d]} q_t(x,y)p(y) \\
&= \sum_{y\in[d]} q_t(x,y)(p(y)-p(x)) \\
&=: q_t(x,\cdot)\cdot \Delta_x p.
\end{aligned}$$

Now, the flow of measures governed by this generator is given by a Fokker-Planck equation:

$$\text{(FP)}\quad \begin{aligned}
\partial_t \mu_t(\{x\}) &= L_t^* \mu_t \\
&= \sum_{y\in E} q_t(y,x)\mu_t(y).
\end{aligned}$$

Now, the Hamiltonian of the MFG is given by

$$H(t, x, \mu, p, \alpha) = [L_t^{\mu,\alpha} p](x) + f(t, x, \mu, \alpha)$$

$$:= \sum_{y \in [d]} \lambda_t(x, y, \mu, \alpha) p(y) + f(t, x, \mu, \alpha).$$

$$= \sum_{y \in [d]} \lambda_t(x, y, \mu, \alpha) \Delta_x p(y) + f(t, x, \mu, \alpha)$$

Taking $H^*(t, x, \mu, p) := \inf_\alpha H(t, x, \mu, p, \alpha)$,

$$\hat{\alpha}(t, x, \mu, p) := \arg\min_\alpha H(t, x, \mu, p, \alpha),$$

which is unique under suitable technical assumptions, we may now characterize the <u>value function</u> of the MFG:

$$u^\mu(t, x) := \inf_\alpha \mathbb{E}\left[\int_t^T f(s, X_s, \mu_s, \alpha_s) ds + g(X_T, \mu_T) \,\middle|\, X_t = x\right].$$

Via standard dynamic programming argument in stochastic control, u^μ solves

$$\begin{cases} \partial_t u^\mu(t, x) + H^*(t, x, \mu_t, \Delta_x u^\mu(t, \cdot)) = 0, & t \in [0, T], x \in [d] \\ u(T, x) = g(x, \mu_T) & \text{(HJB)} \end{cases}$$

<u>Prop:</u> Under suitable assumptions on λ_t, solution to (HJB) exists on $[0, T]$ and is unique, and $\hat{\phi}(t, x) = \hat{\alpha}(t, x, \mu_t, u(t, \cdot))$ is the optimal Markovian control.

How do we find MFG equilibria? As fixed points of a forward-backward system! Specifically, given an initial distribution $\mu_0 = \mathrm{Law}(X_0)$, the flow of measures (μ_t) and value function $u = u^{\mu_t}$ solve:

$$
\text{(FB-ODE)} \quad \begin{cases}
\partial_t \mu_t(x) = \sum_{y \in [d]} \mu_t(y)\, q_t^{\mu,\hat{\alpha}}(y, x), \\
\partial_t u^{\mu_t}(t, x) = -H^*(t, x, \mu_t, \Delta_x u^{\mu_t}(t, \cdot)), \\
\mu_0(x) = \eta(x), \qquad x \in [d], \\
u(T, x) = g(x, \mu_T), \quad x \in [d]
\end{cases}
$$

Moreover, under standard assumptions, an equilibrium with $\mu_t = \mathrm{Law}(X_t)$ exists and is unique.

pf: (Existence sketch)

Take $\mathcal{C}_d = C([0,T]; \mathcal{P}([d]))$, the space of continuous flows of prob. measures on $[d]$. Note that we can continuously embed this into $C([0,T]; \mathbb{R}^d)$. For each $\mu \in \mathcal{C}_d$, take u^μ to be the unique solution to (HJB). From the resulting control, construct

$$
\hat{q}_t^\mu(x, y) = \lambda_t(x, y, \hat{\alpha}(t, x, \mu_t, u^\mu(t, \cdot))),
$$

allowing us to solve (FP) with $\mu_0 = \eta$. This yields a mapping $\Psi : \mathcal{C}_d \to \mathcal{C}_d$; note that solutions to both (HJB) and (FP) are continuous. We now find a fixed point of Ψ.

Now, Ψ is continuous because, under standard assumptions,

$$(\mu, u) \mapsto \hat{\alpha}(t, x, \mu_t, u^{\mu}(t, \cdot))$$

is continuous. Moreover, the rates $\hat{q}_t^{\mu}$ above can be shown to be uniformly bounded, yielding Lipschitz solutions to (FP), with Lipschitz constants <u>independent</u> of μ. In particular, we can invoke the following:

<u>thm:</u> (Schauder fixed point) Let $\mathcal{V}$ be a LCTVS and $E \subseteq \mathcal{V}$ a closed, convex subset of $\mathcal{V}$. If $\Psi : E \to E$ is such that $f(E)$ is rel. compact, then f has a fixed point

Because $\Psi(\mathcal{C}_d)$ is relatively compact by Arzèla-Ascoli, we are done. $\square$

$\mathcal{P}([d]) \hookrightarrow \mathbb{R}^d$ is unif. bounded.

For uniqueness, we require an additional assumption:

<u>Assumption:</u> (Lasry-Lions Monotonicity)

Assume that the terminal cost g is <u>monotone</u> in the sense:

$$\sum_{x \in [d]} [g(x, \mu) - g(x, \mu')] (\mu(x) - \mu'(x)) \geq 0.$$

Moreover, write $f(t, x, \mu, \alpha) = f_0(t, x, \mu) + f_1(t, x, \alpha)$, and assume that $f_0(t, \cdot, \cdot)$ is monotone for $t \in [0, T]$

Uniqueness then follows by considering

optimal controls are unique!

$$J^{\mu}(\hat{\alpha}^{\mu}) - J^{\mu}(\hat{\alpha}^{\mu'}) < 0, \quad J^{\mu'}(\hat{\alpha}^{\mu'}) - J^{\mu'}(\hat{\alpha}^{\mu}) < 0,$$

expanding the cost functionals, and applying monotonicity.

<u>Master Equation</u>

Let us now return to the n-player setting briefly. As before, the n agents have value functions that satisfy

$$v^{n,i}(t,x) = \inf_{\alpha^i} \mathbb{E}\left[\int_t^T f(t, X_s^i, \mu_s^{n,i}, \alpha_s^i)\,ds + g(X_T^i, \mu_T^{n,i}) \;\middle|\; X_t^i = x \right]$$

These satisfy a system of nd^n coupled ODEs, taking the form of (HJB). Specifically, via dynamic programming, this yields a system of HJB equations, one for each agent:

x^i jumps to y

$$0 = \partial_t v^{n,i}(t,x) + \inf_{\alpha_i} \Big\{ f(t, x^i, \mu_t^{n,i}, \alpha^i) + \sum_y \lambda_t(x^i, y, \mu_t^{n,i}, \alpha_i)\left[v^{n,i}(t, x^{-i,y}) - v^{n,i}(t,x) \right] \Big\}$$

another players jumps on inf. fine interval

$$+ \sum_{j \neq i} \sum_y \lambda_t(x^j, y, \mu_t^{n,j}, \hat{\alpha}_t^j)\left[v^{n,i}(t, x^{-j,y}) - v^{n,i}(t,x) \right].$$

Now, we assume by symmetry that $v^{n,i}(t,x) = \mathcal{U}^n(t, x^i, \mu_t^{n,i})$ for some $\mathcal{U}^n : [0,T] \times [d] \times \mathcal{P}([d]) \to \mathbb{R}$. The first part of the Hamiltonian above becomes

$$H^*(t, x, \mu, \Delta_x \mathcal{U}^n(t, \cdot, \mu))$$

almost by definition. What happens when $j \neq i$ jumps to a new state? In this case, we find the decomposition

$$v^{n,i}(t, x^{-j,y}) - v^{n,i}(t,x) = \mathcal{U}^n\left(t, x, \mu + \frac{1}{n-1}(\delta_y - \delta_{x^j})\right) - \mathcal{U}^n(t, x, \mu).$$

Assuming some regularity, we approx. this by

$$\frac{1}{n-1}\left(\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu) - \frac{\partial \mathcal{U}}{\partial \mu(x^j)}(t,x,\mu)\right) + O\left(\frac{1}{n^2}\right)$$

Plugging this in above, we obtain

$$\sum_{j\neq i}\sum_{y} \lambda_t(x^j, y, \mu_t^{n,i}, \hat{\alpha}_t^j)\left[v^{n,i}(t, x^{-j,y}) - v^{n,i}(t,x)\right]$$

$$\approx \sum_{x'\in[d]} (n-1)\mu(x') \sum_{y\in[d]} \lambda_t(x', y, \mu, \hat{\alpha}(t,x,\mu,\mathcal{U}^n))$$
$$\frac{1}{n-1}\left(\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu) - \frac{\partial \mathcal{U}}{\partial \mu(x^j)}(t,x,\mu)\right)$$

$$= \sum_{x'\in[d]}\sum_{y\in[d]} \mu(x')\lambda_t(x',y,\mu,\hat{\alpha}(t,x,\mu,\mathcal{U}^n))\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu)$$
$$\sum_{x'\in[d]}\sum_{y\in[d]} \mu(x')\lambda_t(x',y,\mu,\hat{\alpha})\frac{\partial \mathcal{U}}{\partial \mu(x')}(t,x,\mu)$$

$$= \sum_{x'\in[d]}\sum_{y\in[d]} \mu(x')\lambda_t(x',y,\mu,\hat{\alpha}(t,x,\mu,\mathcal{U}))\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu)$$
$$= \sum_{y\in[d]}\left(\sum_{x'\in[d]} \mu(x')\lambda_t(x',y,\mu,\hat{\alpha}(t,x,\mu,\mathcal{U}))\right)\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu)$$
$$=: \sum_{y\in[d]} h^*(t,\mu,\mathcal{U})(y)\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu).$$

In total, we obtain

$$(\mathrm{ME})\begin{cases} 0 = \partial_t \mathcal{U}(t,x,\mu) + H^*(t,x,\mu,\Delta_x\mathcal{U}(t,\cdot,\mu)) + \\ \qquad \sum_{y\in[d]} h^*(t,\mu,\mathcal{U})(y)\frac{\partial \mathcal{U}}{\partial \mu(y)}(t,x,\mu), \end{cases}$$

with terminal condition $\mathcal{U}(T, x, \mu) = g(x, \mu)$.

Rmk: If $\mathcal{U}: [0,T] \times [d] \times \mathcal{P}([d]) \to \mathbb{R}$ is s.t. $\mathcal{U}(t, x, \cdot)$ is diff. on an open neighborhood of $\mathcal{P}([d]) \hookrightarrow \mathbb{R}^d$, then the above derivative is well-defined. However, noting that

$$\sum_{y \in [d]} h^*(t, \mu, u)(y) = 0,$$

we can rewrite the final term as

$$\sum_{y \neq x} h^*(t, \mu, \mathcal{U}(t, \cdot, \mu))(y) \overbrace{\left(\frac{\partial \mathcal{U}(t, x, \mu)}{\partial \mu(y)} - \frac{\partial \mathcal{U}(t, x, \mu)}{\partial \mu(x)} \right)}^{(*)}.$$

Now, we can:

a) Associate the term $(*)$ with a directional derivative of $\mathcal{U}$ along $\mathcal{P}([d]) \hookrightarrow \mathbb{R}^d$ in the direction $e_y - e_x$, along the simplex.

b) Identify $\mathcal{P}([d])$ with

$$(p_1, \ldots, p_{d-1}) \mapsto (p_1, \ldots, p_{d-1}, 1 - \sum_{i=1}^{d-1} p_i).$$

$$S_{d-1}' := \{ (p_1, \ldots, p_{d-1}) \in [0,1]^{d-1} : \sum_i p_i \leq 1 \},$$

in which case $(*)$ is simply $\frac{\partial \mathcal{U}(t, x, \cdot)}{\partial \mu(y)}$, w.r.t. entries $(\mu(y))_{y \neq x} \in S_{d-1}'$.

In either case, if $\mathcal{U}$ is sufficiently regular, there is no ambiguity concerning spatial derivatives of $\mathcal{U}$ in the master equation!

Now, we have that:

Prop: If $\mathcal{U}$ solves (ME) and μ solves

$$\partial_t \mu_t(x) = h^*(t, \mu_t, \mathcal{U}(t, \cdot, \mu_t))(x)$$

with $\mu_0 = \eta$, then $u(t,x) = \mathcal{U}(t, x, \mu_t)$ solves (HJB), and μ_t solves (KP); (u, μ_t) is then an MFG equilibrium.

pf: We compute

$$\partial_t u(t,x) = \partial_t \mathcal{U}(t, x, \mu_t)$$
$$= \partial_t \mathcal{U}(t, x, \mu_t) + \nabla_\mu \mathcal{U}(t, x, \mu_t) \cdot \partial_t \mu_t$$
$$= -H^*(t, x, \mu_t, \Delta_x \mathcal{U}(t, \cdot, \mu_t))$$
$$= -H^*(t, x, \mu_t, \Delta_x u(t, \cdot)).$$

Moreover,

$$\partial_t \mu_t(x) = \sum_{y \in [d]} \lambda_t(x, y, \mu_t, \hat{\alpha}(t, x, \mu_t, \mathcal{U}(t, \cdot, \mu_t)) \mu_t(y),$$

which is (KP). $\square$

Hence, solving (ME) fully recovers MFG equilibria! In fact, it is much more powerful than just this.

Prop: (CP '18) If H and $\hat{\alpha}$ are locally p-Lipschitz, then the solution $\mathcal{U}$ to (ME) exists uniquely, is C^1 in time, $\nabla_\mu \mathcal{U}(t, x, \cdot)$ is μ-Lipschitz uniformly in time, and if $v^{n,d}$ solves (HJB_n),

$$\sup_{t \in [0,T]} \frac{1}{n} \sum_{i=1}^{n} |v^{n,d}(t,x) - \mathcal{U}(t, x_i, \mu_t^{n,i})| \leq \frac{C}{n},$$

where C is independent of t, x, $\eta = \mathrm{Law}(\mu_0)$, and n.

We also obtain a strong <u>propagation of chaos</u> result via the master equation.

<u>Prop:</u> (CCP '18)

Under the same assumptions, if $Y_i(t)$ are traj. in the n-agent game and $\tilde{X}_i(t)$ $i=1,\dots,n$ are i.i.d. copies of MFG trajectories at equilibrium, then for $i=1,\dots,n$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_i(t)-\tilde{X}_i(t)|\right] \leq C n^{-1/9}$$

for some universal $C>0$.

<u>Rmk:</u> This is akin to a law of large numbers

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\underbrace{\mu_t^n}_{\text{emp. measure}} - \overbrace{\mu_t}^{\text{MFG measure}}|\right] \leq Cn^{-1/9}.$$

Numerical Methods

From the above, we may consider several distinct approaches for computing MFG equilibria:

1) <u>Schauder iteration</u>: simply solve (HJB) and (KP) in an alternating fashion, from an initial guess $(u^{(0)}(t,\cdot), \mu_t^{(0)})$, along fixed time discretization $0=t_0<t_1<\dots<t_M=T$. In practice, this approach can be exceedingly slow or dependent on the initialization.

2) <u>DGME</u>. (CLZ '24): Parametrize solution to (ME) via a neural network, then solve as in physics-informed ML: Writing $\mathcal{L}_t$ as the operator for (ME), minimize

$$L(\theta) := \max_{(t,x,\eta)\in[0,T)\times\mathbb{T}^d\times\mathcal{P}(\mathbb{T}^d)} \left(|\mathcal{L}_t\mathcal{U}(t,x,\eta)| + \right.$$

$$\left| u(T, x, \eta) - g(x, \eta) \right| \Big) .$$

This works decently well for small dimensions, but it is difficult to justify in higher dimensions, and comes equipped with no sample complexity guarantees.

3) Hamiltonian Operator Learning. (HCL '26): if we allow f, g to vary, one may consider a flow map

$$\Phi : (f, g, \eta, t) \longmapsto (u(t, \cdot), \mu_t(\cdot))$$

In my recent work, we show that for parametric families of (f, g), this map is Lipschitz in η and the parameters. Then, from data $\{((f_i, g_i, \eta_i), \Phi(f_i, g_i, \eta_i))\}_{i=1}^{m}$, can learn the flow map to MFG equilibria, complete with sample + parametric complexity guarantees!