v-topology on $\mathcal{P}_v(\mathsf{X})$. Using the dual formulation of W_p [42, Theorem 5.10], we can write β_v as

$$\beta_v(\mu,\nu) = \sup_{\substack{(h,g)\in\mathcal{L}_1(\mu)\times\mathcal{L}_1(\nu):\\ h(x)-g(y)\le d_{\mathsf{X}}(x,y)^p}} |\mu(h)-\nu(g)|^{\frac{1}{p}} + |\mu(v)-\nu(v)|, \tag{2.3}$$

where $\mathcal{L}_1(\lambda)$ denotes the set of all λ-integrable real functions on X.

Remark 1. In the remainder of the paper, $\mathcal{P}(\mathsf{X})$ is always equipped with the weak topology while $\mathcal{P}_v(\mathsf{X})$ is always equipped with the v-topology. In other words, when we say that a function over $\mathcal{P}_v(\mathsf{X})$ is continuous, it should be understood that it is continuous with respect to v-topology. Similarly, a function over $\mathcal{P}(\mathsf{X})$ is continuous if it is continuous with respect to weak topology.

Assumption 1.

(a) The one-stage cost function c is continuous. ✓

(b) A is compact. ✓

(c) There exists a nonnegative real number α such that

moment function

any continuous

$$\sup_{(a,\mu)\in\mathsf{A}\times\mathcal{P}(\mathsf{X})} \int_{\mathsf{X}} w(y)p(dy|x,a,\mu) \le \alpha w(x).$$

In our case, the state transition is deterministic; so it's automatically true

(d) The stochastic kernel $p(\,\cdot\,|x,a,\mu)$ is weakly continuous; that is, if $(x_n,a_n,\mu_n)\to(x,a,\mu)$ in $\mathsf{X}\times\mathsf{A}\times\mathcal{P}(\mathsf{X})$, then $p(\,\cdot\,|x_n,a_n,\mu_n)\to p(\,\cdot\,|x,a,\mu)$ weakly. In addition, the function $\int_{\mathsf{X}} w(y)p(dy|x,a,\mu)$ is continuous in (x,a,μ). ✓

(e) The initial probability measure μ_0 satisfies

defined in 2.1

$$\int_{\mathsf{X}} v(x)\mu_0(dx) =: M < \infty. \quad ✓$$

cost function

Remark 2. Note that Assumption 1(c) implies that the range of p lies in $\mathcal{P}_v(\mathsf{X})$; that is, $\{p(\,\cdot\,|x,a,\mu) : (x,a,\mu)\in\mathsf{X}\times\mathsf{A}\times\mathcal{P}(\mathsf{X})\}\subset\mathcal{P}_v(\mathsf{X})$. Therefore, Assumption 1(d) is equivalent to the following condition: if $(x_n,a_n,\mu_n)\to(x,a,\mu)$ in $\mathsf{X}\times\mathsf{A}\times\mathcal{P}(\mathsf{X})$, then $p(\,\cdot\,|x_n,a_n,\mu_n)\to p(\,\cdot\,|x,a,\mu)$ with respect to the v-topology on $\mathcal{P}_v(\mathsf{X})$.

For each $t\ge 0$, let us define

$$\mathcal{P}_v^t(\mathsf{X}) := \left\{\mu\in\mathcal{P}_v(\mathsf{X}) : \int_{\mathsf{X}} w(x)\mu(dx)\le\alpha^t M\right\}.$$

(f) There exist $\gamma\ge 1$ and $R\ge 0$ such that, for each $t\ge 0$, why?

$$\sup_{(a,\mu)\in\mathsf{A}\times\mathcal{P}_v^t(\mathsf{X})} c(x,a,\mu)\le M_t v(x), \qquad \text{where } M_t := \gamma^t R.$$

If c is bdd, say $|c(x,a,\mu)|\le R$, then can choose $\gamma=1$. Then (g) is automatically satisfied w/ a discount factor

(g) We assume that $\alpha\beta\gamma<1$. discount factor

Assumption 2.

Define the following moduli of continuity:

$$\omega_p(r) := \sup_{(x,a)\in\mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu:\\ \tilde{\rho}_v(\mu,\nu)\le r}} \|p(\,\cdot\,|x,a,\mu)-p(\,\cdot\,|x,a,\nu)\|_v,$$

$$\omega_c(r) := \sup_{(x,a)\in\mathsf{X}\times\mathsf{A}} \sup_{\substack{\mu,\nu:\\ \tilde{\rho}_v(\mu,\nu)\le r}} |c(x,a,\mu)-c(x,a,\nu)|,$$

where $\tilde{\rho}_v = \beta_v$ (see (2.3)) if c is unbounded, and $\tilde{\rho}_v = \rho$ if c is bounded. For any function $g : \mathcal{P}_v(\mathsf{X}) \to \mathbb{R}$, we define the v-norm of g as follows:

$$\|g\|_v^* := \sup_{\mu \in \mathcal{P}_v(\mathsf{X})} \frac{|g(\mu)|}{\mu(v)}.$$

(h) We assume that $\omega_p(r) \to 0$ and $\omega_c(r) \to 0$ as $r \to 0$. Moreover, for any $\mu \in \mathcal{P}_v(\mathsf{X})$, $\|\omega_p(\tilde{\rho}_v(\,\cdot\,,\mu))\|_v^* < \infty$ and $\|\omega_c(\tilde{\rho}_v(\,\cdot\,,\mu))\|_v^* < \infty$.

(i) There exists a nonnegative real number B such that

$$\sup_{(a,\mu) \in \mathsf{A} \times \mathcal{P}_v(\mathsf{X})} \int_{\mathsf{X}} v^2(y) p(dy|x,a,\mu) \leq B v^2(x).$$

Remark 3. Note that, if the state transition probability p is independent of the mean-field term, then Assumption 2(h) for ω_p is always true. Indeed, in that case, we have $\omega_p(r) = 0$ for all r.

Remark 4. Suppose that $v = w$. Define a metric λ_{X} on X as follows:

$$\lambda_{\mathsf{X}}(x,y) := \begin{cases} 0 & \text{if } x = y, \\ v(x) + v(y) & \text{if } x \neq y. \end{cases}$$

For any real-valued measurable function g on X, define the Lipschitz seminorm of g:

$$\|g\|_\lambda := \sup_{x \neq y} \frac{|g(x) - g(y)|}{\lambda_{\mathsf{X}}(x,y)}.$$

Let $\mathrm{Lip}_\lambda(1,\mathbb{R}) := \{g : \|g\|_\lambda \leq 1\}$. Then, for any $\mu, \nu \in \mathcal{P}_v(\mathsf{X})$, we have

$$\|\mu - \nu\|_v = \sup_{g \in \mathrm{Lip}_\lambda(1,\mathbb{R})} \left| \int_{\mathsf{X}} g(x)\mu(dx) - \int_{\mathsf{X}} g(x)\nu(dx) \right| \tag{2.4}$$

[17, Lemma 2.1]. This alternative formulation of v-norm will be useful when verifying Assumption 2(h) for specific examples.

Remark 5. In the remainder of this paper, all the proofs are obtained under the assumption that the cost function c is unbounded. The bounded case can be covered by slight modification of the proofs for the unbounded case.

Before proceeding to the next section, we state an important (but straightforward) result which will be used in what follows. It basically states that there is no loss of generality in restricting the infima in Definition 2.2 to weakly continuous Markov policies.

THEOREM 2.3. *Under Assumption* 1, *for any policy* $\pi^{(N)} \in \mathsf{M}^{(N)}$ *we have*

$$\inf_{\pi^i \in \mathsf{M}_i} J_i^{(N)}(\pi_{-i}^{(N)}, \pi^i) = \inf_{\pi^i \in \mathsf{M}_i^c} J_i^{(N)}(\pi_{-i}^{(N)}, \pi^i) \qquad \forall i = 1, \ldots, N.$$

Proof. See [40, Appendix A]. □

3. Mean-field games and mean-field equilibria. We begin by considering a mean-field game that can be interpreted as the infinite-population limit $N \to \infty$ of the game introduced in the preceding section. This mean-field game is specified by the quintuple $(\mathsf{X}, \mathsf{A}, p, c, \mu_0)$, where, as before, X and A denote the state and action