

SEQUENTIAL GAMES

Rajesh Dachiraju
Hyderabad, India
rajesh.dachiraju@gmail.com

ABSTRACT. We formulate sequential games in which players take a sequence of decisions in each turn, and players may randomize over HMMs. Payoffs are assumed to be BV for these games. We prove an existence theorem for the Nash equilibrium.

1. INTRODUCTION

Sequential games model decisions whose strategic effects unfold over multiple stages. In this paper, each player chooses a hidden Markov model (HMM) that generates a finite sequence of probability vectors. A completion rule combines players' sequences, including when their lengths differ, and the payoff is the sum of stage rewards over the completed game. The HMM therefore represents a probability law over a player's pure sequences and serves as that player's mixed strategy. For existence, the induced payoff functions on strategy profiles are required to be BV and essentially bounded. Our main question is when such games admit a Nash equilibrium. We give a finite-horizon existence theorem using Kakutani's fixed-point theorem.

2. LITERATURE SURVEY

Nash's work provides the classical starting point for non-cooperative equilibrium. In a finite strategic-form game, each player has a finite pure-strategy set and may randomize over it; Nash proves the existence of an equilibrium in mixed strategies [5]. The present paper changes the strategy objects: a pure strategy is a sequence of probability vectors, while an HMM specifies a probability law over such sequences.

For games with more general strategy spaces, Glicksberg extended fixed-point methods to compact strategy spaces and continuous payoff functions [2].

A related dynamic framework is Shapley's theory of stochastic games, in which play proceeds through states and stages, with transitions between states and rewards accumulated over time [7]. That work supplies an important precedent for games with temporal evolution. The present formulation is different in its choice of strategy representation: each player selects an HMM law over an entire variable-length sequence, and the payoff

is assigned to the completed sequence profile. The extensive-form tradition provides another central foundation for sequential interaction. Kuhn’s treatment of extensive games formalizes the order of decisions and the information available to players at their decision points [4]. Kreps and Wilson later introduced sequential equilibrium, requiring sequential rationality together with consistent beliefs in extensive games with imperfect information [3]. These frameworks focus on how actions and information unfold within a game tree. The present model instead takes each player’s complete finite sequence of probability vectors as the output of an HMM and evaluates the resulting completed profile; it does not, as currently specified, model strategies contingent on the observed history of play.

Repeated-game theory studies strategic interaction across multiple plays of a stage game. For example, Fudenberg and Maskin characterize equilibrium payoff possibilities in infinitely repeated games under discounting or incomplete information [1]. That literature highlights how continuation incentives and history-dependent strategies can shape outcomes. The time-series games considered here have finite outputs almost surely and use an undiscounted sum of stage rewards over the completed game. Thus the repeated-game results provide useful context, but do not directly establish the equilibrium claim for the HMM strategy class studied here.

HMMs are established probabilistic models for sequential observations. Rabiner’s tutorial describes their basic formulation and inference problems [6]. Here the HMM is used as a player’s mixed strategy, so the game-theoretic question concerns equilibrium among HMM-generated distributions rather than statistical inference alone.

Taken together, these references motivate the ingredients of the present model—mixed strategies, fixed-point existence methods, stage-based dynamic games, and HMM sequence models—while leaving the specific combination studied here to be analyzed under explicit payoff-regularity and HMM density assumptions.

3. SEQUENTIAL GAMES

3.1. Pure strategies. Let m denote the number of available static actions. Write

$$\Delta_m = \left\{ x \in [0, 1]^m : \sum_{a=1}^m x_a = 1 \right\}$$

for the probability simplex over these actions. A pure strategy of player i is a finite sequence

$$s_i = (x_{i,1}, \dots, x_{i,T_i}), \quad x_{i,q} \in \Delta_m,$$

where the length $T_i \geq 1$ may vary across strategies. Thus each stage of a pure strategy specifies a probability vector over the player’s static actions. Infinite sequences are not included in the pure-strategy space.

3.2. Mixed strategies. A mixed strategy for player i may be an admissible hidden Markov model (HMM), which generates a random pure strategy s_i , or a probability mixture over admissible HMMs. Under a mixture, the player first draws an HMM according to a probability distribution over the admissible HMM family and then draws that HMM's output sequence. Thus an HMM or mixture over HMMs induces a probability law on player i 's set of finite sequences. We require every admissible HMM to terminate almost surely. The players randomize independently.

3.3. Payoff. For each player, the payoff rule assigns a real number to each completed game. Let n be the number of players. At stage q , collect the players' probability vectors into a matrix $A_q \in \mathbb{R}^{m \times n}$, with the i -th column representing player i 's vector. The feasible stage-profile space is

$$\Gamma = \left\{ A \in [0, 1]^{m \times n} : \sum_{a=1}^m A_{ai} = 1 \text{ for every } i = 1, \dots, n \right\}.$$

When players' sequence lengths differ, a specified completion rule determines stage profiles after a player's output ends. For a pure strategy profile $s = (s_1, \dots, s_n)$, let $\Phi(s) = (A_1, \dots, A_T)$ be the resulting completed game, where $T = \max_i T_i$. The game space Ω consists of these finite sequences, with different lengths treated as distinct games.

For each player i , let $\zeta_i : \Gamma \rightarrow \mathbb{R}$ be the stage-reward function. The payoff rule on completed games is

$$g_i : \Omega \rightarrow \mathbb{R}, \quad g_i(\omega) = \sum_{q=1}^T \zeta_i(A_q), \quad \omega = (A_1, \dots, A_T).$$

The actual payoff assigned to a pure strategy profile is the induced function

$$p_i(s) := (g_i \circ \Phi)(s) = g_i(\Phi(s)).$$

Thus p_i , rather than the payoff rule g_i , is the payoff function used in strategic analysis and in the equilibrium theorem's regularity condition.

Every realized game has finite length, so the realized reward sequence is square-summable whenever each stage reward is real-valued. With bounded stage rewards and a common finite horizon, each completed-game payoff is measurable and bounded.

As the length of every realized game is finite, it can be easily seen that if the stage reward payoff functions ζ_i are BV then the induced actual payoff functions p_i are also BV. Thus for these games the functions ζ_i are assumed to be BV and hence the actual induced payoffs p_i are BV.

We establish sufficient conditions for Nash equilibrium in sequential games. All games are assumed to be non-co-operative.

4. AN EXISTENCE THEOREM FOR SEQUENTIAL GAMES

We give a sufficient condition for the existence of a Nash equilibrium in a time-series game whose players choose hidden Markov models (HMMs)

generating finite, variable-length pure strategies. Each admissible HMM terminates after finitely many stages with probability one. Each pure strategy is a finite sequence of probability vectors, and the payoff of a complete game is the sum of its stage rewards, which may be signed. The result uses a fixed-point argument on compact convex sets of normalized HMM densities. Bounded variation of individual payoffs and densities is not, by itself, enough for this argument: compactness of the feasible density sets and control of output lengths are also required.

4.1. Model. Let $n \geq 2$ be the number of players and m the number of static actions. Define the closed probability simplex

$$\Delta_m = \left\{ x \in [0, 1]^m : \sum_{a=1}^m x_a = 1 \right\}.$$

A pure strategy of player i is a finite sequence of vectors in Δ_m ; infinite sequences are excluded. In the existence theorem, we impose a common finite bound L on the output length. Thus the theorem uses the pure-strategy space

$$S_i = \bigcup_{T=1}^L \Delta_m^T.$$

The finite union is understood as a disjoint union, so sequences of different lengths are distinct strategies. For the density formulation, equip each component Δ_m^T with its normalized Hausdorff measure λ_T on its affine hull and use the finite direct-sum reference measure $\mu_i = L^{-1} \bigoplus_{T=1}^L \lambda_T$. Every HMM law represented by a density is required to be absolutely continuous with respect to μ_i . For a profile $s = (s_1, \dots, s_n)$, a specified measurable rule combines the players' sequences into the completed game $\Phi(s) \in \Omega$. The rule must specify how the game is represented when the players' output lengths differ; the game ends when the longest output ends.

For each player i , let $\zeta_i : \Gamma \rightarrow \mathbb{R}$ be measurable and bounded. The payoff for a completed game $\omega = (A_1, \dots, A_T)$ is

$$g_i(\omega) = \sum_{q=1}^T \zeta_i(A_q).$$

With bounded stage rewards and $T \leq L$, each completed-game payoff is measurable and bounded, and square-summability holds automatically. The actual payoff of a strategy profile is the induced function $p_i(s) = g_i(\Phi(s))$; the equilibrium theorem imposes BV regularity and essential boundedness on this function. For finite sequences whose lengths have no common deterministic bound, square-summability alone does not ensure integrability of the total payoff; a model without a common bound must additionally require integrability under every admissible HMM profile.

A mixed strategy for player i is a probability density f_i over S_i , with respect to a fixed reference measure μ_i . It is normalized:

$$f_i \geq 0, \quad \int_{S_i} f_i d\mu_i = 1.$$

The players' HMM outputs are independent. The expected payoff at a profile $f = (f_1, \dots, f_n)$ is

$$U_i(f) = \int_{S_1 \times \dots \times S_n} p_i(s_1, \dots, s_n) \prod_{j=1}^n f_j(s_j) d(\mu_1 \times \dots \times \mu_n).$$

A profile f^* is a Nash equilibrium if

$$U_i(f^*) \geq U_i(h_i, f_{-i}^*)$$

for every player i and every admissible alternative density h_i .

4.2. Assumptions for existence. For each player i , define $\mathcal{H}_i \subset L^1(\mu_i)$ as the admissible strategy-density set, allowing arbitrary probability mixtures over HMM-generated laws and their L^1 limits. The following assumptions are established from the HMM conditions in the next subsection:

- (1) $\mathcal{H}_i$ is nonempty, convex, and compact in the $L^1(\mu_i)$ norm.
- (2) Every $f_i \in \mathcal{H}_i$ is a normalized probability density:

$$f_i \geq 0 \quad \mu_i\text{-almost everywhere}, \quad \int_{S_i} f_i d\mu_i = 1.$$

- (3) The game-completion map $\Phi : S_1 \times \dots \times S_n \rightarrow \Omega$ is measurable.

4.3. Nonemptiness, compactness, and convexity from HMM mixtures. For each player i , suppose the family of continuous symbol-emission densities is nonempty and compact in L^1 , with a common finite BV bound. Fix a finite hidden-state space and a common deterministic output-length bound L . Let the admissible initial and transition parameters form a nonempty closed subset of their finite-dimensional probability simplices. These parameter sets are compact; no BV regularity is required of transition probabilities.

Let Θ_i be the resulting compact HMM parameter space, and let $Q_i(\theta) \in L^1(\mu_i)$ denote the output-sequence density generated by $\theta \in \Theta_i$. For each length profile, $Q_i(\theta)$ is a finite sum over hidden-state paths of products of normalized emission densities, weighted by initial and transition probabilities. The product and finite-mixture operations are continuous in L^1 , so $\theta \mapsto Q_i(\theta)$ is continuous. Therefore

$$K_i = \{Q_i(\theta) : \theta \in \Theta_i\}$$

is a nonempty compact subset of $L^1(\mu_i)$, and every element of K_i is a normalized probability density.

Define the admissible strategy-density set by allowing arbitrary probability mixtures over the HMM laws and their L^1 limits:

$$\mathcal{H}_i = \overline{\text{co}}(K_i) \subset L^1(\mu_i).$$

It is nonempty because K_i is nonempty, and convex by construction. The closed convex hull of a compact subset of a Banach space is compact, hence $\mathcal{H}_i$ is compact in $L^1(\mu_i)$. Nonnegativity and unit integral are preserved under convex combinations and L^1 limits, so every element of $\mathcal{H}_i$ is a normalized density. This establishes Assumptions 1–2; the measurable completion rule in the model establishes Assumption 3.

4.4. Generator bounds in a continuous-time extension. A bounded generator controls regularity of the hidden-state marginal trajectory, but this is only one part of the HMM density-set problem. For a continuous-time Markov chain with a fixed number K of hidden states, let $p(t)$ be a row probability vector satisfying

$$\dot{p}(t) = p(t)Q,$$

where Q is a generator (its off-diagonal entries are nonnegative and each row sums to zero). If

$$\|Q\|_\infty := \max_{1 \leq r \leq K} \sum_{s=1}^K |Q_{rs}| \leq M,$$

then

$$\|\dot{p}(t)\|_1 = \|p(t)Q\|_1 \leq \|p(t)\|_1 \|Q\|_\infty \leq M.$$

Consequently,

$$\text{Var}_{[0,T]}(p; \ell^1) \leq \int_0^T \|\dot{p}(t)\|_1 dt \leq MT.$$

Thus a uniform generator bound gives a uniform BV bound (indeed, a uniform Lipschitz bound) on the hidden-state marginal paths. For fixed K , initial distributions in the simplex, and generators in a closed bounded set, these paths form a compact family in $C([0, T], \mathbb{R}^K)$.

4.5. Existence theorem.

Theorem 4.1. *Under the results established in Sections 4.1 and 4.3, Assumptions 1–3 hold. If for every player the induced payoff p_i is BV and essentially bounded on each finite-horizon profile component, there exists a Nash equilibrium $f^* \in \prod_{i=1}^n \mathcal{H}_i$.*

Proof. Let

$$X = \prod_{i=1}^n \mathcal{H}_i, \quad p_i(s_1, \dots, s_n) = g_i(\Phi(s_1, \dots, s_n)).$$

By the payoff regularity condition in the theorem, p_i is measurable and essentially bounded on each component; we use a bounded representative in the payoff integral. For $f = (f_1, \dots, f_n) \in X$, write its joint density as

$$F_f(s_1, \dots, s_n) = \prod_{j=1}^n f_j(s_j).$$

The expected payoff

$$U_i(f) = \int p_i(s) F_f(s) d(\mu_1 \times \dots \times \mu_n)(s)$$

is therefore well-defined and finite.

For any two profiles $f, g \in X$, expand the difference of their product densities as

$$F_f - F_g = \sum_{k=1}^n \left(\prod_{j < k} f_j \right) (f_k - g_k) \left(\prod_{j > k} g_j \right).$$

All component densities are normalized. Integrating the absolute value of each summand and using Fubini's theorem gives

$$\|F_f - F_g\|_{L^1(\mu_1 \times \dots \times \mu_n)} \leq \sum_{j=1}^n \|f_j - g_j\|_{L^1(\mu_j)}.$$

Consequently,

$$|U_i(f) - U_i(g)| \leq \|p_i\|_\infty \sum_{j=1}^n \|f_j - g_j\|_{L^1(\mu_j)}.$$

Thus U_i is jointly continuous on X . For fixed f_{-i} , it is affine in f_i , since $F_{(f_i, f_{-i})}$ is linear in f_i .

For each player i , define the best-response correspondence

$$B_i(f_{-i}) = \arg \max_{h_i \in \mathcal{H}_i} U_i(h_i, f_{-i}).$$

Because $\mathcal{H}_i$ is nonempty and compact and U_i is continuous, the maximum is attained. Hence $B_i(f_{-i})$ is nonempty and compact. Affineness in h_i and convexity of $\mathcal{H}_i$ imply that $B_i(f_{-i})$ is convex.

The graph of B_i is closed: if $f_{-i}^k \rightarrow f_{-i}$, $h_i^k \rightarrow h_i$, and $h_i^k \in B_i(f_{-i}^k)$, then for every $r_i \in \mathcal{H}_i$,

$$U_i(h_i^k, f_{-i}^k) \geq U_i(r_i, f_{-i}^k).$$

Taking limits and using continuity yields $U_i(h_i, f_{-i}) \geq U_i(r_i, f_{-i})$, so $h_i \in B_i(f_{-i})$. Since the domain and range are compact metric spaces, the closed-graph property implies that B_i is upper hemicontinuous. Now define $B : X \rightrightarrows X$ by

$$B(f) = \prod_{i=1}^n B_i(f_{-i}).$$

This correspondence has nonempty, compact, convex values and is upper hemicontinuous. The set X is nonempty, compact, and convex by assumptions 1–2. Kakutani’s fixed-point theorem therefore gives $f^* \in X$ such that $f^* \in B(f^*)$. By the definition of each best-response set, no player can improve their expected payoff by unilaterally changing their admissible strategy density. Thus f^* is a Nash equilibrium. $\square$

4.6. Scope of the result. Sections 4.1 and 4.3 establish the measurable completion rule and compact convex strategy-density sets under the stated finite-horizon and symbol-emission conditions. The theorem then requires BV regularity and essential boundedness of each induced payoff on the finite-horizon strategy-profile space. In particular, the finite-horizon formulation assumes a common deterministic length bound L . Almost-sure termination alone is weaker: it permits arbitrarily long outputs and does not give a compact finite-horizon strategy space. If output lengths are unbounded, a different existence argument must control the tails uniformly and ensure that expected payoffs are well defined and continuous under changes in strategy.

For a finite-state HMM with a strictly left-to-right transition structure and no self-transitions, the number of emitted stages is bounded by the number of states, provided the model’s termination convention permits only forward moves and termination. This is one way to obtain a common bound L when the state count is uniformly bounded.

A compact family of symbol-emission densities with a common BV bound, together with a fixed finite state space and finite output horizon, yields a compact family of HMM sequence densities. Allowing probability mixtures over those HMM laws gives compact convex strategy-density sets.

An ideal estimation method does not, on its own, replace these existence hypotheses. Suppose that, as the sample size grows, an estimator consistently recovers every relevant HMM output density and parameter. This is a statistical statement about recovering a data-generating model (subject to the estimator’s assumptions); Nash equilibrium is a strategic statement requiring a profile at which no player can gain by deviating. Consistent estimation does not ensure that such a profile exists, that best responses attain their suprema, or that limits of estimated strategies remain admissible. Compactness, convexity, and payoff continuity address these issues in the fixed-point proof.

The distinction also matters computationally. If equilibrium existence is established by another argument, and estimated payoffs and deviations approximate the true ones uniformly over the relevant strategy sets, consistent density estimation may help construct approximate equilibria. Pointwise convergence of estimated HMM densities or parameters alone does not provide that uniform control and does not establish an exact equilibrium. An ideal estimator may remove a practical obstacle to learning or computing strategies, but it cannot by itself replace the mathematical existence argument or its hypotheses. The theorem requires each induced payoff on the

finite-horizon strategy-profile space to be BV and essentially bounded. Measurability of the completion rule then makes the expected-payoff expression well defined, and the L^1 estimate in the proof supplies continuity in the players' densities. A compact symbol-emission density family with a common BV bound, finite state space, finite horizon, and mixtures over HMM laws supplies the compact convex density sets required by the theorem.

5. POTENTIAL PRACTICAL APPLICATIONS

The framework can represent finite interactions in which each player commits to an HMM that generates a finite sequence of probability vectors. At each stage, the players' vectors form the columns of a stage-profile matrix, and a specified stage-reward function assigns each player a reward. The total payoff is the sum of those rewards over the completed game, including any stages supplied by the completion rule when output lengths differ. An HMM's hidden states describe the internal evolution of the player's randomized sequence; they do not, without additional modelling, represent observed opponent actions or guarantee adaptation to public history.

- **Capacity and resource allocation.** Players such as service providers or infrastructure operators may generate sequences of probability vectors over allocation options. Stage rewards can encode delivered service, operating costs, congestion, and unmet demand. A model must specify how the allocation vectors jointly determine outcomes and how play proceeds after a player's sequence ends.
- **Communication and network access.** Competing devices or providers may generate probability vectors over channels, routes, or access options. The stage-reward functions can represent throughput, delay, interference, and resource use. Latent HMM states can encode a player's internal operating modes; if external congestion is to affect decisions, the model must specify how that information enters the player's strategy.
- **Market participation.** Traders or institutions may generate sequences of probability vectors over market actions or asset choices. Stage rewards can account for returns, transaction costs, and risk exposure, provided the market outcome and reward are defined as measurable functions of the stage-profile matrices. The finite-horizon payoff assumptions must be checked for the chosen reward rule.
- **Security and defense.** An attacker and a defender may generate probability vectors over targets, monitoring choices, or protective actions. Stage rewards can combine losses, defensive costs, and successful interventions. A model intended to capture learning from observed moves must add an explicit information and response mechanism; autonomous HMM state evolution alone does not encode that feature.

- **Inventory and supply chains.** Suppliers, carriers, and retailers may generate sequences of probability vectors over production, replenishment, or shipping options. Stage rewards can include fulfillment, holding costs, shortages, and waste. The completion rule should specify the operational meaning of stages after an agent's output ends.

Across these examples, the dimension m is the number of entries in each player's stage probability vector, and (n) is the number of players, as in the model. Applying the existence theorem requires a concrete measurable completion rule, BV payoffs, and induced profile-space payoffs that are BV and essentially bounded. Sections 4.1 and 4.3 provide the measurable completion rule and compact convex strategy-density sets when the common output-length bound and symbol-emission conditions hold.

REFERENCES

1. Drew Fudenberg and Eric Maskin, *The folk theorem in repeated games with discounting or with incomplete information*, *Econometrica* **54** (1986), no. 3, 533–554.
2. Irving L. Glicksberg, *A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points*, *Proceedings of the American Mathematical Society* **3** (1952), no. 1, 170–174.
3. David M. Kreps and Robert Wilson, *Sequential equilibria*, *Econometrica* **50** (1982), no. 4, 863–894.
4. Harold W. Kuhn, *Extensive games and the problem of information*, *Contributions to the Theory of Games* (Harold W. Kuhn and Albert W. Tucker, eds.), *Annals of Mathematics Studies*, vol. 2, Princeton University Press, 1953, pp. 193–216.
5. John Nash, *Non-cooperative games*, *Annals of Mathematics* **54** (1951), no. 2, 286–295.
6. Lawrence R. Rabiner, *A tutorial on hidden markov models and selected applications in speech recognition*, *Proceedings of the IEEE* **77** (1989), no. 2, 257–286.
7. Lloyd S. Shapley, *Stochastic games*, *Proceedings of the National Academy of Sciences* **39** (1953), no. 10, 1095–1100.