

Exponential convergence of fictitious-play FBSDEs for finite-player stochastic differential games

Kristoffer Andersson

University of Verona

`kristofferherbert.andersson@univr.it`

Joint work with Adam Andersson and Per Ljung

Vienna Seminar in Mathematical Finance and Probability

January 29, 2026

Table of contents

- 1 Game formulation
- 2 HJB characterization of Nash equilibrium
- 3 Nash and fictitious-play FBSDEs
 - Nash FBSDE
 - Fictitious-play FBSDEs
 - Numerically tractable surrogate FBSDEs
- 4 Exponential convergence of fictitious-play FBSDEs
 - Theoretical results
 - Empirical results
- 5 Full approximation of the Nash FBSDE

Game formulation

Stochastic differential game: Domain and controls

- We study an N -player stochastic differential game driven by a k -dimensional Brownian motion.
- Players are indexed by $\mathcal{I} = \{1, \dots, N\}$.
- The state of the game evolves in a (possibly unbounded) spatial domain $\mathcal{D} \subset \mathbb{R}^n$, with its closure denoted by $\overline{\mathcal{D}}$.
- The parabolic domain is

$$Q_T = [0, T) \times \mathcal{D},$$

with parabolic boundary

$$\partial Q_T = ([0, T) \times \partial \mathcal{D}) \cup (\{T\} \times \overline{\mathcal{D}}).$$

Stochastic differential game: State dynamics and exit time

- Each player i uses a **Markovian feedback control** $\alpha^i \in \mathbb{A}^i$, where $\mathbb{A}^i$ denotes a class of admissible functions such that, for all $(t, \mathbf{x}) \in Q_T$,

$$\alpha^i(t, \mathbf{x}) \in A^i \subset \mathbb{R}^d,$$

and each action set A^i is compact.

- The joint control is

$$\alpha(t, \mathbf{x}) = (\alpha^1(t, \mathbf{x}), \dots, \alpha^N(t, \mathbf{x})).$$

- The controlled state process $\mathbf{X}^{t, \mathbf{x}, \alpha} : [t, T] \times \Omega \rightarrow \overline{\mathcal{D}}$ is the predictable, square-integrable process satisfying, for all $s \in [t, T]$,

$$\mathbf{X}_s^{t, \mathbf{x}, \alpha} = \mathbf{x} + \int_t^{s \wedge \tau} \bar{\mathbf{b}}(r, \mathbf{X}_r^{t, \mathbf{x}, \alpha}, \alpha(r, \mathbf{X}_r^{t, \mathbf{x}, \alpha})) \, dr + \int_t^{s \wedge \tau} \Sigma(r, \mathbf{X}_r^{t, \mathbf{x}, \alpha}) \, dW_r.$$

- The stopping time τ is the **exit time** of the parabolic domain Q_T ,

$$\tau = \tau(t, \mathbf{x}, \alpha) = \inf\{s \geq t : (s, \mathbf{X}_s^{t, \mathbf{x}, \alpha}) \in \partial Q_T\}.$$

- In the special case $\mathcal{D} = \mathbb{R}^n$, the parabolic boundary reduces to $\{T\} \times \mathbb{R}^n$, hence $\tau = T$ almost surely.

Cost functionals and Nash equilibrium

- For an initial state $(t, \mathbf{x}) \in Q_T$, each player $i \in \mathcal{I}$ seeks to minimize the cost functional

$$J^i(t, \mathbf{x}; \alpha) = \mathbb{E}_{t, \mathbf{x}} \left[\int_t^\tau \bar{f}^i(s, \mathbf{X}_s^{t, \mathbf{x}, \alpha}, \alpha(s, \mathbf{X}_s^{t, \mathbf{x}, \alpha})) ds + g^i(\mathbf{X}_\tau^{t, \mathbf{x}, \alpha}) \right],$$

where τ is the exit time from Q_T .

- For a joint action $\mathbf{a} = (a^1, \dots, a^N)$ and $i \in \mathcal{I}$, we write $\mathbf{a}^{-i} = (a^1, \dots, a^{i-1}, a^{i+1}, \dots, a^N)$, and for $b^i \in A^i$ define $[b^i, \mathbf{a}^{-i}]_i = (a^1, \dots, a^{i-1}, b^i, a^{i+1}, \dots, a^N)$. We use the same convention for $\mathbb{A}$ and $\mathbb{A}^{-i}$.
- A **closed-loop Nash equilibrium** is a policy tuple $\alpha^* = (\alpha^{1,*}, \dots, \alpha^{N,*})$ such that no player can decrease their cost by unilaterally deviating.
- Formally, for every $i \in \mathcal{I}$, every alternative control $\alpha^i \in \mathbb{A}^i$, and every $(t, \mathbf{x}) \in Q_T$,

$$J^i(t, \mathbf{x}; \alpha^*) \leq J^i(t, \mathbf{x}; [\alpha^i, \alpha^{-i,*}]_i).$$

HJB characterization of Nash equilibrium

Best-response problem: value functions and HJB equations

- Given the policies of the other players α^{-i} , player i faces a single-player stochastic control problem on Q_T .
- The **value function** is

$$V^i(t, \mathbf{x}; \alpha^{-i}) = \inf_{\alpha^i \in \mathbb{A}^i} J^i(t, \mathbf{x}; [\alpha^i, \alpha^{-i}]_i), \quad (t, \mathbf{x}) \in Q_T.$$

- Associated with the best-response problem of player i is the HJB equation on Q_T .

$$\partial_t \mathcal{V}^i + \mathcal{L}_t \mathcal{V}^i + H^i(t, \mathbf{x}, D_{\mathbf{x}} \mathcal{V}^i(t, \mathbf{x}; \alpha^{-i}), \alpha^{-i}(t, \mathbf{x})) = 0, \quad (t, \mathbf{x}) \in Q_T,$$

with boundary condition

$$\mathcal{V}^i(t, \mathbf{x}; \alpha^{-i}) = g^i(t, \mathbf{x}), \quad (t, \mathbf{x}) \in \partial Q_T.$$

- Here, the Hamiltonian $H^i: Q_T \times \mathbb{R}^n \times A^{-i} \rightarrow \mathbb{R}$, for $(t, \mathbf{x}, p, \mathbf{a}^{-i}) \in Q_T \times \mathbb{R}^n \times A^{-i}$ is given by

$$H^i(t, \mathbf{x}, p, \mathbf{a}^{-i}) := \inf_{\mathbf{a}^i \in A^i} \left(\langle \bar{\mathbf{b}}(t, \mathbf{x}, [\mathbf{a}^i, \mathbf{a}^{-i}]_i), p \rangle + \bar{f}^i(t, \mathbf{x}, [\mathbf{a}^i, \mathbf{a}^{-i}]_i) \right),$$

Verification of best-response policies

- Under suitable regularity assumptions, a verification theorem implies that the solution of the HJB equation coincides with the value function, i.e. for $(t, \mathbf{x}, \alpha^{-i}) \in Q_T \times \mathbb{A}^{-i}$,

$$V^i(t, \mathbf{x}, \alpha^{-i}) = \mathcal{V}^i(t, \mathbf{x}, \alpha^{-i}).$$

- moreover, the optimal policy $\beta^i: Q_T \times \mathbb{A}^{-i} \rightarrow A^i$, for $(t, \mathbf{x}, \alpha^{-i}) \in Q_T \times \mathbb{A}^{-i}$ is given by

$$\beta^i(t, \mathbf{x}, \alpha^{-i}) = \kappa^i(t, \mathbf{x}, D_{\mathbf{x}} \mathcal{V}^i(t, \mathbf{x}, \alpha^{-i}), \alpha^{-i}(t, \mathbf{x})).$$

Here, $(t, \mathbf{x}) \mapsto \beta^i(t, \mathbf{x}, \alpha^{-i})$ in $\mathbb{A}^i$ is the policy (given α^{-i}) and $\kappa^i: Q_T \times \mathbb{R}^n \times A^{-i} \rightarrow A^i$, for $(t, \mathbf{x}, p, \mathbf{a}^{-i}) \in Q_T \times \mathbb{R}^n \times A^{-i}$ is given by

$$\kappa^i(t, \mathbf{x}, p, \mathbf{a}^{-i}) \in \arg \min_{\mathbf{a}^i \in A^i} \left(\langle \bar{\mathbf{b}}(t, \mathbf{x}, [\mathbf{a}^i, \mathbf{a}^{-i}]_i), p \rangle + \bar{f}^i(t, \mathbf{x}, [\mathbf{a}^i, \mathbf{a}^{-i}]_i) \right).$$

- We assume that $\kappa^1, \dots, \kappa^N$ exist and are unique, i.e., for every $(t, \mathbf{x}, p, \mathbf{a}^{-i}) \in Q_T \times \mathbb{R}^n \times A^{-i}$ the associated minimization problem admits a unique minimizer.

Vectorized best-responses and fixed-point formulation

- For $(t, \mathbf{x}, \mathbf{a}, \alpha) \in Q_T \times A \times \mathbb{A}$ and $\mathbf{p} = (p^1, \dots, p^N) \in \mathbb{R}^{n \times N}$, define the vectorized maps by

$$\begin{aligned}\mathbf{V}(t, \mathbf{x}, \alpha) &= (V^1(t, \mathbf{x}, \alpha^{-1}), \dots, V^N(t, \mathbf{x}, \alpha^{-N})), \\ \kappa(t, \mathbf{x}, \mathbf{p}, \mathbf{a}) &= (\kappa^1(t, \mathbf{x}, p^1, \mathbf{a}^{-1}), \dots, \kappa^N(t, \mathbf{x}, p^N, \mathbf{a}^{-N})), \\ \beta(t, \mathbf{x}, \alpha) &= \kappa(t, \mathbf{x}, D_{\mathbf{x}} \mathbf{V}(t, \mathbf{x}, \alpha), \alpha(t, \mathbf{x})).\end{aligned}$$

- A Nash equilibrium is characterized as a fixed point

$$\alpha^* = \beta(\alpha^*), \quad \beta: \mathbb{A} \rightarrow \mathbb{A}.$$

- For $(t, \mathbf{x}) \in Q_T$, the equilibrium control $\alpha_p^*(t, \mathbf{x}) \in A$ solves

$$\alpha_p^*(t, \mathbf{x}) = \kappa(t, \mathbf{x}, \mathbf{p}, \alpha_p^*(t, \mathbf{x})).$$

- Assume that the **generalized Isaacs condition** holds, i.e. for $\mathbf{p} \in \mathbb{R}^{n \times N}$, the fixed-point problem admits a solution $\alpha_p^* \in \mathbb{A}$, and the map

$$\mathbf{c}(t, \mathbf{x}, \mathbf{p}) := \alpha_p^*(t, \mathbf{x})$$

is well defined from $Q_T \times \mathbb{R}^{n \times N}$ to A .

- Under the generalized Isaacs condition, setting $\mathbf{p} = D_{\mathbf{x}} \mathbf{V}(t, \mathbf{x})$, and assuming sufficiently regular solutions of the Nash HJB system yields a closed-loop Nash equilibrium.

Nash and fictitious-play FBSDEs

Nash FBSDE formulation I

- The Nash HJB system admits an equivalent FBSDE representation on $[0, \tau]$, where τ is the exit time from Q_T .
- The forward equation is

$$\mathbf{X}_t = \mathbf{x} + \int_0^{t \wedge \tau} \bar{\mathbf{b}}(s, \mathbf{X}_s, \mathbf{c}(s, \mathbf{X}_s, D_{\mathbf{x}} \mathbf{V}(t, \mathbf{X}_t))) ds + \int_0^{t \wedge \tau} \Sigma(s, \mathbf{X}_s) d\mathbf{W}_s.$$

- The backward equation is

$$\mathbf{Y}_t = g(\tau, \mathbf{X}_\tau) + \int_t^\tau \bar{\mathbf{f}}(s, \mathbf{X}_s, \mathbf{c}(s, \mathbf{X}_s, D_{\mathbf{x}} \mathbf{V}(t, \mathbf{X}_t))) ds - \int_t^\tau \mathbf{Z}_s^\top d\mathbf{W}_s.$$

- Since Σ is uniformly parabolic, $\Sigma(t, \mathbf{x})\Sigma^\top(t, \mathbf{x})$ is invertible for $(t, \mathbf{x}) \in Q_T$ and we denote by $\Phi: Q_T \rightarrow \mathbb{R}^{n \times k}$, defined by $(t, \mathbf{x}) \mapsto (\Sigma(t, \mathbf{x})\Sigma^\top(t, \mathbf{x}))^{-1}\Sigma(t, \mathbf{x})$

$$D_{\mathbf{x}} \mathbf{V}(t, \mathbf{X}_t) = \Phi(t, \mathbf{X}_t) \mathbf{Z}_t.$$

Hence the equilibrium control is

$$\alpha^*(t, \mathbf{X}_t) = \mathbf{c}(t, \mathbf{X}_t, \Phi(t, \mathbf{X}_t) \mathbf{Z}_t).$$

- The Markovian representation of the backward components is

$$\mathbf{Y}_t = \mathbf{V}(t, \mathbf{X}_t), \quad \mathbf{Z}_t = \Sigma^\top(t, \mathbf{X}_t) D_{\mathbf{x}} \mathbf{V}(t, \mathbf{X}_t).$$

Nash FBSDE formulation II

By inserting the feedback control $\mathbf{c}(t, \mathbf{x}, D_{\mathbf{x}}V)$ from the Nash HJB system and using the Markovian relation $Z^i = \Sigma^\top D_{\mathbf{x}} \mathbf{V}^i$, we rewrite $\bar{\mathbf{b}}$ and $\bar{\mathbf{f}}$ as functions of $(t, \mathbf{X}_t, Z_t^i, \mathbf{Z}_t^{-i})$, yielding the FBSDE system below.

- The forward equation is

$$\mathbf{X}_t^i = \mathbf{x} + \int_0^{t \wedge \tau} \mathbf{b}^i(s, \mathbf{X}_s^i, Z_s^i, \mathbf{Z}_s^{-i}) ds + \int_0^{t \wedge \tau} \Sigma(s, \mathbf{X}_s^i) d\mathbf{W}_s.$$

- The backward equation is

$$Y_t^i = g(\tau, \mathbf{X}_\tau^i) + \int_t^\tau f^i(s, \mathbf{X}_s^i, Z_s^i, \mathbf{Z}_s^{-i}) ds - \int_t^\tau (Z_s^i)^\top d\mathbf{W}_s.$$

- The Markovian representation is

$$Y_t^i = V^i(t, \mathbf{X}_t^i), \quad Z_t^i = \Sigma^\top(t, \mathbf{X}_t^i) D_{\mathbf{x}} V^i(t, \mathbf{X}_t^i).$$

Note: all players share the same forward state ($\mathbf{X}^i = \mathbf{X}^j$), but their backward components differ because each player has its own value function. This notation is convenient when fictitious play is introduced. From now on, we therefore write $\mathbf{X}$ for the common forward state in the Nash FBSDE.

Standing assumptions

General assumptions (both bounded and unbounded domains):

- Coefficients b^i , Σ , f^i are Lipschitz in $\mathbf{x}$ and Hölder-1/2 in t .
- Uniform parabolicity: there exists $c, C > 0$ with $c < C$ such that for all $(t, \mathbf{x}) \in Q_T$ and $\xi \in \mathbb{R}^n$ we have $c\|\xi\|^2 \leq \xi^\top \Sigma(t, \mathbf{x}) \Sigma^\top(t, \mathbf{x}) \xi \leq C\|\xi\|^2$.
- The feedback map $c(\cdot, \mathbf{x}, p)$ is Lipschitz in $(\mathbf{x}, p)$.
- For each player $i \in \mathcal{I}$, the action set A^i is assumed to be compact.

Additional assumptions for bounded domains $|\mathcal{D}| < \infty$:

- The spatial domain $\mathcal{D}$ has a $C^{2+\alpha}$ boundary, for $\alpha \in (0, 1)$.
- The parabolic boundary data satisfy $g^i \in C^{1+\alpha/2, 2+\alpha}(\partial Q_T)$, ensuring compatibility with the interior $C^{1+\alpha/2, 2+\alpha}$ regularity of the HJB system.
- For $\mathbf{p} \in \mathbb{R}^n$, $H^i(\cdot, \cdot, p)$, $\Sigma \in C^{\alpha/2, \alpha}(Q_T)$,
- The drift and diffusion coefficients b^i and Σ are continuously differentiable with bounded derivatives in $\mathbf{x}$ and p for b^i and $\mathbf{x}$ for Σ .

(Additional minor technical assumptions appear in the full analysis, but the structural conditions are those listed above.)

Fictitious-play FBSDEs

- We start from an initial collection of Markovian Z -maps

$$\zeta^{(0)} = (\zeta^{1,(0)}, \dots, \zeta^{N,(0)}), \quad \zeta^{i,(0)} : Q_T \rightarrow \mathbb{R}^k.$$

- At iteration $m \geq 1$, each player $i \in \mathcal{I}$ computes a best response to the opponents' maps $\zeta^{-i,(m-1)}$ by solving the stopped FBSDE (denoted $\text{FBSDE}(i, m)$):

$$\begin{aligned} \mathbf{X}_t^{i,(m)} &= \mathbf{x} + \int_0^{t \wedge \tau^{i,(m)}} b^i(s, \mathbf{X}_s^{i,(m)}, Z_s^{i,(m)}, \zeta^{-i,(m-1)}(s, \mathbf{X}_s^{i,(m)})) ds \\ &\quad + \int_0^{t \wedge \tau^{i,(m)}} \Sigma(s, \mathbf{X}_s^{i,(m)}) dW_s, \end{aligned}$$

$$\begin{aligned} Y_t^{i,(m)} &= g^i(\tau^{i,(m)}, \mathbf{X}_{\tau^{i,(m)}}^{i,(m)}) + \int_t^{\tau^{i,(m)}} f^i(s, \mathbf{X}_s^{i,(m)}, Z_s^{i,(m)}, \zeta^{-i,(m-1)}(s, \mathbf{X}_s^{i,(m)})) ds \\ &\quad - \int_t^{\tau^{i,(m)}} (Z_s^{i,(m)})^\top dW_s. \end{aligned}$$

- The Z -process admits a Markovian representation:

$$\zeta^{i,(m)}(t, \mathbf{X}_t^{i,(m)}) = Z_t^{i,(m)}, \quad t \leq \tau^{i,(m)}.$$

- Thus iteration m produces a new collection of Z -maps.

Surrogate FBSDE

- In practice, player i does not have access to the exact opponents' feedback maps $\zeta^{-i,(m-1)}(t, \mathbf{x})$. Instead we work with approximate feedback maps

$$\tilde{\zeta}^{-i,(m-1)}(t, \mathbf{x}).$$

- The **surrogate FBSDE** for player i at iteration m (denoted $\widetilde{\text{FBSDE}}(i, m)$) is the stopped system

$$\begin{aligned} \tilde{\mathbf{X}}_t^{i,(m)} &= \mathbf{x} + \int_0^{t \wedge \tilde{\tau}^{i,(m)}} b^i\left(s, \tilde{\mathbf{X}}_s^{i,(m)}, \tilde{\mathbf{Z}}_s^{i,(m)}, \tilde{\zeta}^{-i,(m-1)}(s, \tilde{\mathbf{X}}_s^{i,(m)})\right) ds \\ &\quad + \int_0^{t \wedge \tilde{\tau}^{i,(m)}} \Sigma(s, \tilde{\mathbf{X}}_s^{i,(m)}) dW_s, \end{aligned}$$

$$\begin{aligned} \tilde{\mathbf{Y}}_t^{i,(m)} &= g^i(\tilde{\tau}^{i,(m)}, \tilde{\mathbf{X}}_{\tilde{\tau}^{i,(m)}}^{i,(m)}) + \int_t^{\tilde{\tau}^{i,(m)}} f^i\left(s, \tilde{\mathbf{X}}_s^{i,(m)}, \tilde{\mathbf{Z}}_s^{i,(m)}, \tilde{\zeta}^{-i,(m-1)}(s, \tilde{\mathbf{X}}_s^{i,(m)})\right) ds \\ &\quad - \int_t^{\tilde{\tau}^{i,(m)}} (\tilde{\mathbf{Z}}_s^{i,(m)})^\top dW_s. \end{aligned}$$

- This is the time-continuous FBSDE we approximate numerically at each fictitious-play iteration.

Approximation hierarchy

Hierarchical structure of the approximation pipeline

$$\widetilde{\text{FBSDE}}^{\pi}(i, m) \longrightarrow \widetilde{\text{FBSDE}}(i, m) \longrightarrow \text{FBSDE}(i, m) \longrightarrow \text{Nash FBSDE}$$

Today's focus: Exponential convergence in m of

$$\text{FBSDE}(i, m) \longrightarrow \text{Nash FBSDE}.$$

A similar argument establishes convergence of the surrogate system $\widetilde{\text{FBSDE}}(i, m)$ to $\text{FBSDE}(i, m)$.

Exponential convergence of fictitious-play FBSDEs

Notation for the main result

- **FBSDE solutions.** For each player $i \in \mathcal{I}$,

$(\mathbf{X}, Y^i, Z^i)$ solves the Nash FBSDE, $(\mathbf{X}^{i,(m)}, Y^{i,(m)}, Z^{i,(m)})$ solves FBSDE(i, m).

- **Forward and backward errors.**

$$\delta \mathbf{X}_t^{i,(m)} := \mathbf{X}_t - \mathbf{X}_t^{i,(m)}, \quad \delta Y_t^{i,(m)} := Y_t^i - Y_t^{i,(m)}.$$

- **Strategy (functional) error.**

$$\delta \bar{\zeta}_t^{i,(m)} := \zeta^i(t, \mathbf{X}_t) - \zeta^{i,(m)}(t, \mathbf{X}_t).$$

This is the quantity whose norm must contract from one fictitious-play iteration to the next.

- **Stopped processes.** For any adapted process A and stopping time τ ,

$$A_t^{[\tau]} := A_t \mathbf{1}_{\{t \leq \tau\}}.$$

Main result (bounded domain)

Under the standing assumptions for the bounded-domain case ($|\mathcal{D}| < \infty$), the fictitious-play FBSDEs converge exponentially fast to the Nash FBSDE.

Theorem (Exponential convergence of fictitious-play FBSDEs)

Assume $|\mathcal{D}| < \infty$. Then for every $K \in (0, 1)$ there exists a constant $C < \infty$ (depending only on $N, T, \mathcal{D}, M, K, \mathbf{x}, b, \mathbf{f}, \Sigma, g$) such that

$$\max_{i \in \mathcal{I}} \|\delta \mathbf{X}^{i,(m)}\|_{\mathbb{S}_{0,T}^{2,n}}^2 + \sum_{i \in \mathcal{I}} \|\delta \mathbf{Y}^{i,(m)}\|_{\mathbb{S}_{0,T}^{2,1}}^2 + \sum_{i \in \mathcal{I}} \left\| \zeta^{i,[T]}(\cdot, \mathbf{X}_\cdot) - \zeta^{i,(m),[\tau^{i,(m)}]}(\cdot, \mathbf{X}_\cdot^{i,(m)}) \right\|_{\mathbb{H}_{0,T}^{2,k}}^2 \leq CK^m.$$

Remark (unbounded domain $\mathcal{D} = \mathbb{R}^n$). In the unbounded-domain case, the same exponential convergence result holds provided that either

- the Lipschitz constants are sufficiently small, or
- the time horizon T is taken sufficiently small.

Lemma 1: Forward error estimate

Lemma (Informal forward error bounds)

For each $i \in \mathcal{I}$ and iteration $m \geq 1$, there exists a constant $C < \infty$, such that

$$\|\delta \mathbf{X}^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{S}_{0,T}^{2,n}}^2 \leq C \left(\|\delta \bar{\zeta}^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 + \sum_{j \in \mathcal{I} \setminus \{i\}} \|\delta \bar{\zeta}^{j,(m-1),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 \right),$$

and on the full time interval,

$$\|\delta \mathbf{X}^{i,(m)}\|_{\mathbb{S}_{0,T}^{2,n}}^2 \leq C \left(\mathbb{E}|\delta \tau^{i,(m)}| + \|\delta \bar{\zeta}^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 + \sum_{j \in \mathcal{I} \setminus \{i\}} \|\delta \bar{\zeta}^{j,(m-1),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 \right),$$

where $\tau_{\min}^{i,(m)} := \tau \wedge \tau^{i,(m)}$.

Interpretation. Two bounds are needed. The error up to the first exit time is controlled purely by the functional strategy errors via Lipschitz continuity, BDG, and Grönwall.

The full error includes contributions from the interval $[\tau_{\min}^{i,(m)}, \tau_{\max}^{i,(m)}]$, producing a term $\mathbb{E}|\delta \tau^{i,(m)}|$ that cannot be absorbed back. Hence the stopping-time error must be handled separately. The disjoint-support decomposition enables this split; afterwards all remaining estimates are routine SDE arguments.

Lemma 2: Backward error estimate

Lemma (Informal backward error bounds)

For each $i \in \mathcal{I}$ and $m \geq 1$, there exists a constant $C < \infty$, such that

$$\|\delta Y^{i,(m)}\|_{\mathbb{S}_{0,T}^{2,1}}^2 \leq C \left(\mathbb{E}|\delta \tau^{i,(m)}| + \|\delta \zeta^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 + \sum_{j \in \mathcal{I} \setminus \{i\}} \|\delta \zeta^{j,(m-1),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 \right),$$

and

$$\begin{aligned} & \|\zeta^{i,[\tau]}(\cdot, \mathbf{X}_{\cdot}) - \zeta^{i,(m),[\tau_{\min}^{i,(m)}]}(\cdot, \mathbf{X}_{\cdot}^{i,(m)})\|_{\mathbb{H}_{0,T}^{2,k}}^2 \\ & \leq C \left(\mathbb{E}|\delta \tau^{i,(m)}| + \|\delta \zeta^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 + \sum_{j \in \mathcal{I} \setminus \{i\}} \|\delta \zeta^{j,(m-1),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{H}_{0,T}^{2,k}}^2 \right). \end{aligned}$$

Interpretation. This is a standard BSDE stability estimate: apply Itô to $|\delta Y|^2$ and use BDG, Young, Cauchy–Schwarz, and Lipschitz continuity, together with bounds from Lemma 1.

The only nonstandard feature is that the two BSDEs are stopped at different exit times. A disjoint-support decomposition separates terms before and after the first stopping time. After the split, the remaining estimates are routine.

Remaining tasks

The bounds in Lemmas 1–2 leave two remaining error terms to control:

1. Stopping-time error.

- We also need to bound $\mathbb{E}|\delta\tau^{i,(m)}|$, which feeds into Lemmas 1–2.
- Standard exit-time estimates from parabolic PDE theory relate this to the forward error up the first stopping time, i.e., $\mathbb{E}|\delta\tau^{i,(m)}| \leq C\|\delta\mathbf{X}^{i,(m),[\tau_{\min}^{i,(m)}]}\|_{\mathbb{S}_{0,T}^{2,n}}$. (See e.g.¹)

2. Functional strategy error.

- We must control $\delta\bar{\zeta}^{i,(m),[\tau_{\min}^{i,(m)}]}$, $i \in \mathcal{I}$.
- The functional error is the same object for all players, but each term is stopped at its own time $\tau_{\min}^{i,(m)} = \tau \wedge \tau^{i,(m)}$.
- In a contraction argument:
 - LHS contains a single norm (stopped at $\tau_{\min}^{i,(m)}$),
 - RHS contains N versions of the same error, each stopped at a different random time.
- Hence the norms do not match and the misaligned stopping times prevent a direct contraction.

¹Dokuchaev, Nikolai. *Estimates for distances between first exit times via parabolic equations in unbounded cylinders*. Probability theory and related fields 129.2 (2004): 290-314.

Lemma 3: Contraction of functional strategy errors

Lemma (Exponential contraction of functional strategy errors)

For each $K \in (0, 1)$ there exists a constant $C < \infty$ (depending only on $N, T, \mathcal{D}, M, K, \mathbf{x}, b, f, \Sigma, g$) such that

$$\sum_{i \in \mathcal{I}} \|\delta \bar{\zeta}^{i, (m), [\tau]}\|_{\mathbb{H}_{0, T}^{2, k}}^2 \leq CK^m, \quad m \geq 1.$$

Interpretation. The functional errors of all players, stopped at the true exit time τ , decay exponentially fast with the fictitious-play iteration m . Since $\tau \geq \tau_{\min}^{i, (m)}$ for all i , this bound also controls the functional errors on the shorter intervals $[0, \tau_{\min}^{i, (m)}]$ appearing in Lemmas 1–2. Combined with the forward and backward error estimates, this contraction drives all components of the fictitious-play FBSDEs towards the Nash FBSDE.

Caveat. Deriving this contraction relies on gradient bounds for the associated parabolic PDEs: for each m and i we need a classical solution $V^{i, (m)}$ with a spatial gradient $\nabla_{\mathbf{x}} V^{i, (m)}$ that is uniformly bounded in m .

Sketch of proof (Step 1): The auxiliary FBSDE

Auxiliary FBSDE. Applying Itô's formula to $V^{i,(m)}(t, \mathbf{X}_t)$ along the true state process $\mathbf{X}$ yields the system

$$\begin{aligned} \mathbf{X}_t &= \mathbf{x} + \int_0^{t \wedge \tau} b^i(s, \mathbf{X}_s, \zeta^i(s, \mathbf{X}_s), \zeta^{-i}(s, \mathbf{X}_s)) ds + \int_0^{t \wedge \tau} \Sigma(s, \mathbf{X}_s) d\mathbf{W}_s, \\ Y_t^{i,(m),\mathbf{X}} &= g^i(\tau, \mathbf{X}_\tau) + \int_{t \wedge \tau}^{\tau} \left[f^i(s, \mathbf{X}_s, \zeta^{i,(m)}(s, \mathbf{X}_s), \zeta^{-i,(m-1)}(s, \mathbf{X}_s)) \right. \\ &\quad \left. + \langle \delta b_s^{i,(m)}, D_{\mathbf{X}} V^{i,(m)}(s, \mathbf{X}_s) \rangle \right] ds - \int_{t \wedge \tau}^{\tau} (\zeta^{i,(m)}(s, \mathbf{X}_s))^\top d\mathbf{W}_s, \end{aligned}$$

where for $t \in [0, \tau]$,

$$\delta b_t^{i,(m)} := b^i(t, \mathbf{X}_t, \zeta^i(t, \mathbf{X}_t), \zeta^{-i}(t, \mathbf{X}_t)) - b^i(t, \mathbf{X}_t, \zeta^{i,(m)}(t, \mathbf{X}_t), \zeta^{-i,(m-1)}(t, \mathbf{X}_t)).$$

Key point: The control keeps the same functional form $\zeta^{i,(m)}(t, \mathbf{X}_t)$ as in FBSDE(i, m), but now everything is valid up to the common exit time τ .

Sketch of proof (Step 2): Bound the functional strategy errors

Applying Itô's lemma to $e^{\beta t} |Y_t^i - Y_t^{i,(m),\mathbf{X}}|^2$ and using standard inequalities yields, for any $\beta > 0$,

$$\left\| \delta \bar{\zeta}^{i,(m),[\tau]} \right\|_{\mathbb{H}_{\beta,0,T}^{2,k}}^2 \leq 2\beta^{-1} \left(\left\| \delta f^{i,(m),[\tau]} \right\|_{\mathbb{H}_{\beta,0,T}^{2,1}}^2 + \left\| \langle \delta b^{i,(m)}, D_x V^{i,(m)}(\cdot, \mathbf{X}) \rangle \right\|_{\mathbb{H}_{\beta,0,T}^{2,1}}^2 \right).$$

Here $\|\cdot\|_{\mathbb{H}_{\beta,0,T}^{2,k}}$ is the weighted $\mathbb{H}^2$ -norm.

For the second term, Cauchy–Schwarz and a uniform spatial gradient bound give

$$\left\| \langle \delta b^{i,(m)}, D_x V^{i,(m)}(\cdot, \mathbf{X}) \rangle \right\|_{\mathbb{H}_{\beta,0,T}^{2,1}}^2 \leq M \left\| \delta b^{i,(m),[\tau]} \right\|_{\mathbb{H}_{\beta,0,T}^{2,n}}^2,$$

with

$$M := \sup_{m \geq 1} \max_{i \in \mathcal{I}} \sup_{(t,x) \in [0,T] \times \mathcal{D}} \|D_x V^{i,(m)}(t, x)\|_{\mathbb{R}^n}^2.$$

For now, assume $M < \infty$.

Sketch of proof (Step 3): Collecting terms to obtain contraction

Using the Lipschitz continuity of f^i and b^i , and summing the estimate from Step 2 over $i \in \mathcal{I}$, we obtain

$$\sum_{i \in \mathcal{I}} \|\delta \bar{\zeta}^{i,(m),[\tau]}\|_{\mathbb{H}_{\beta,0,T}^{2,k}}^2 \leq K \sum_{i \in \mathcal{I}} \|\delta \bar{\zeta}^{i,(m-1),[\tau]}\|_{\mathbb{H}_{\beta,0,T}^{2,k}}^2,$$

where

$$K := \frac{2(L_\ell + L_b M)(N-1)}{\beta - 2(L_\ell + L_b M)}.$$

Since we may choose β sufficiently large so that $K \in (0, 1)$, this inequality is contractive in m . Because the weighted and unweighted $\mathbb{H}^2$ -norms are equivalent, the same contraction holds in $\mathbb{H}_{0,T}^{2,k}$.

Remaining task: establish that $M < \infty$.

Sketch of proof (Final step): Gradient bound

Theorem (Parabolic Schauder estimate)

For some $\alpha \in (0, 1)$, assume $\partial\mathcal{D} \in C^{2+\alpha}$, $\Sigma, h(\cdot, \cdot, p) \in C^{\alpha/2, \alpha}(Q_T)$, $g \in C^{1+\alpha/2, 2+\alpha}(Q_T)$, Σ uniformly parabolic on Q_T , and

$$|h(t, x, p) - h(t, x, \tilde{p})| \leq L_h \|p - \tilde{p}\|, \quad \|h(\cdot, \cdot, p)\|_{C^{\alpha/2, \alpha}(Q_T)} \leq C_h(1 + \|p\|).$$

Then the PDE admits a unique solution $u \in C^{1+\alpha/2, 2+\alpha}(Q_T)$.

Consequence. Applying this to each $V^{i, (m)}$ yields a uniform gradient bound

$$M := \sup_{m \geq 1} \max_{i \in \mathcal{I}} \sup_{(t, x) \in [0, T] \times \mathcal{D}} \|D_x V^{i, (m)}(t, x)\|_{\mathbb{R}^n} < \infty,$$

which completes the contraction argument.

Remark. A uniform bound is obtained because the dependence on $\zeta^{-i, (m)}$ in the coefficients remains controlled when the action sets A^i are compact.

Bounded vs. unbounded domain

Unbounded domain $\mathcal{D} = \mathbb{R}^n$.

- The issue with different stopping times disappears (all processes stop at T), so a direct kick-back argument yields a contraction.
- To ensure the contraction constant lies in $(0, 1)$, one needs either small Lipschitz coefficients or a small terminal time T .
- The terminal condition requires less regularity, since no compatibility with a spatial boundary is needed.

Bounded domain $|\mathcal{D}| < \infty$.

- Exit times lead to player-specific stopping times $\tau_{\min}^{i,(m)}$, and the contraction argument is substantially more delicate.
- A uniform contraction requires *parabolic PDE regularity*, including smooth boundary $\partial\mathcal{D}$ and compatible terminal data.
- No smallness assumptions on Lipschitz constants or T are needed.

Returning to the approximation hierarchy

Three levels of convergence

$$\widetilde{\text{FBSDE}}^\pi(i, m) \xrightarrow{\text{(order 1/2, unbounded domain)}} \widetilde{\text{FBSDE}}(i, m)$$

$$\widetilde{\text{FBSDE}}(i, m) \xrightarrow{\text{(exponential in } m)}} \text{FBSDE}(i, m)$$

$$\text{FBSDE}(i, m) \xrightarrow{\text{(exponential in } m)}} \text{Nash FBSDE}$$

Remarks.

- The exponential convergence $\text{FBSDE}(i, m) \rightarrow \text{Nash FBSDE}$ was the main focus of this talk.
- The surrogate system exhibits the same exponential rate.
- The discretized scheme converges with order 1/2 in unbounded domains. In bounded domains, stopping times reduce the Euler–Maruyama forward rate to order 1/4, and a full deep BSDE error theory is still missing.

Inter-bank lending game (linear–quadratic model)

- The state $X_t = (X_t^1, \dots, X_t^N) \in \mathbb{R}^N$ represents the log-reserves of N banks, with empirical mean

$$\bar{X}_t := \frac{1}{N} \sum_{j=1}^N X_t^j.$$

- Each bank $i \in \mathcal{I}$ controls its drift through α^i :

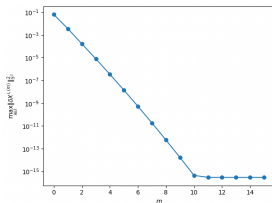
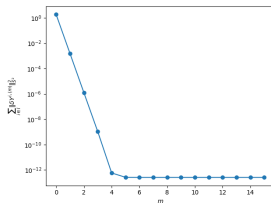
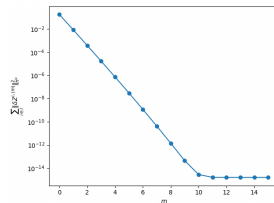
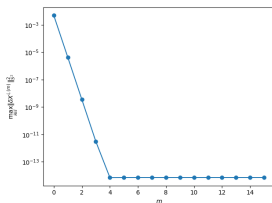
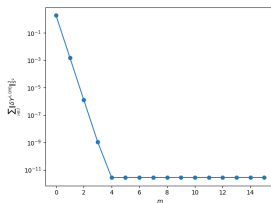
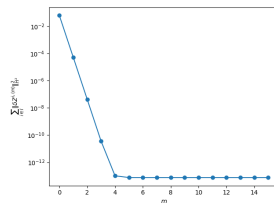
$$X_t^i = X_0^i + \int_0^t \left(a(\bar{X}_s - X_s^i) + \alpha_s^i \right) ds + \sigma \widehat{W}_t^i, \quad \widehat{W}_t^i = \rho W_t^0 + \sqrt{1 - \rho^2} W_t^i.$$

- The cost functional of bank i is

$$J^i(\alpha) = \mathbb{E} \left[\int_0^T \left(\frac{1}{2} (\alpha_t^i)^2 - q \alpha_t^i (\bar{X}_t - X_t^i) + \frac{\varepsilon}{2} (\bar{X}_t - X_t^i)^2 \right) dt + \frac{c}{2} (\bar{X}_T - X_T^i)^2 \right].$$

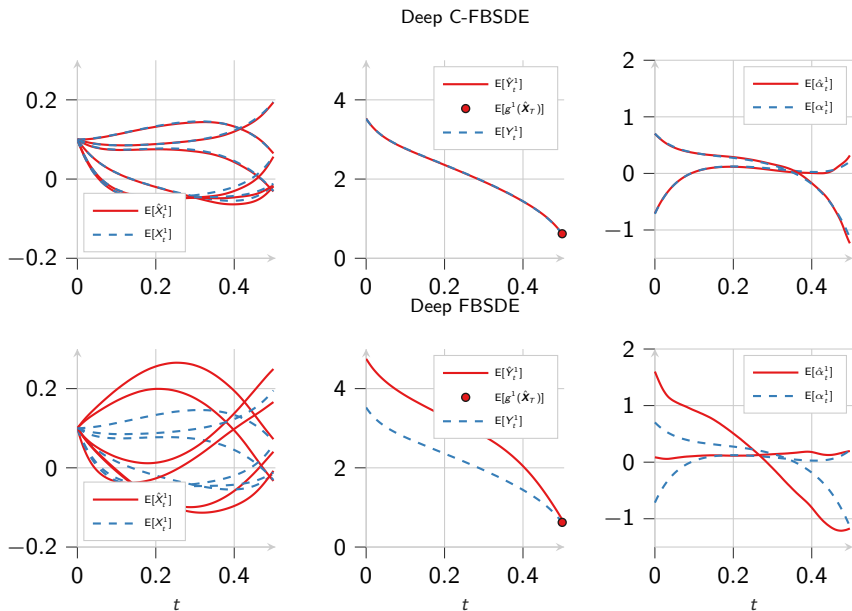
- This is a symmetric linear–quadratic stochastic differential game with an explicit closed-loop Nash equilibrium.

Exponential convergence of fictitious play (empirical results)

(a) $N = 2, X$ (b) $N = 2, Y$ (c) $N = 2, Z$ (d) $N = 20, X$ (e) $N = 20, Y$ (f) $N = 20, Z$

Full approximation of the Nash FBSDE

Three-player game with 12-dimensional state space (Player 1)



Contact Information

Email address

`kristofferherbert.andersson@univr.it`

Related Publications

Han, Jiequn, Hu, Ruimeng, and Long, Jihao. *Convergence of deep fictitious play for stochastic differential games*. arXiv preprint arXiv:2008.05519 (2020).

Andersson, Kristoffer, Andersson, Adam, and Cornelis, W. Oosterlee. *The Deep Multi-FBSDE Method: A Robust Deep Learning Method for Coupled FBSDEs*. arXiv preprint arXiv:2503.13193. (2025).

Andersson, Kristoffer, Andersson, Adam, and Cornelis, W. Oosterlee. *Convergence of a Robust Deep FBSDE Method for Stochastic Control*. *SIAM Journal on Scientific Computing* 45, (1), (2023): A226–A255.

Back-up slides

Jacobi vs. Gauss–Seidel updates in fictitious play (extra slide)

- **Jacobi update (parallel)** All players compute their best responses *simultaneously* using the previous iteration's maps:

$$\zeta^{(m)} = \left(\zeta^{1,(m)}(\cdot; \zeta^{-(m-1)}), \dots, \zeta^{N,(m)}(\cdot; \zeta^{-(m-1)}) \right).$$

Uses only $\zeta^{(m-1)}$ as input.

Advantages:

- Can compute all FBSDEs for players $i = 1, \dots, N$ **in parallel**.
- natural for large N (multi-agent setting).
- **Gauss–Seidel update (sequential)** Players update one at a time, using the *most recently updated* maps of the others:

$$\zeta^{(m)} = \left(\zeta^{1,(m)}(\cdot; \zeta^{-(m-1)}), \zeta^{2,(m)}(\cdot; \zeta^{1,(m)}, \zeta^{3,(m-1)}, \dots), \dots \right).$$

Disadvantages:

- Updates must be computed **sequentially** (no parallelization).
- Harder to analyze for fully coupled FBSDE systems.
- More expensive for large N .
- **Why we use Jacobi:** It allows us to solve all player-specific FBSDEs in parallel at each iteration m , which is essential for efficiency in multi-player stochastic differential games.

Continuous and discrete stochastic control problems

Time continuous problem:

$$\begin{cases} X_t^\zeta = x_0 + \int_0^t \bar{b}(s, X_s^\zeta, \zeta(s, X_s^\zeta)) ds + \int_0^t \sigma(s, X_s^\zeta) dW_s, \\ J(\zeta) = \mathbb{E} \left[\int_0^T \bar{f}(t, X_t^\zeta, u(t, X_t^\zeta)) dt + g(X_T^\zeta) \right]. \end{cases}$$

Time discrete problem:

$$\begin{cases} X_{n+1}^{\zeta_h, h} = X_n^{\zeta_h, h} + \bar{b}(t_n, X_n^{\zeta_h, h}, \zeta_{h,n}(X_n^{\zeta_h, h})) h + \sigma(t_n, X_n^{\zeta_h, h}) \Delta W_n, \\ J_h(\zeta_h) = \mathbb{E} \left[g(X_N^{\zeta_h, h}) + \sum_{n=0}^{N-1} \bar{f}(t_n, X_n^{\zeta_h, h}, \zeta_{h,n}(X_n^{\zeta_h, h})) h \right], \end{cases}$$

Optimal Markov maps:

$$\zeta^* \in \arg \min_{\zeta} J(\zeta), \quad \zeta_h^* \in \arg \min_{\zeta_h} J_h(\zeta_h).$$

Sufficient consistency assumption to relax small T assumption in Theorem 1:

$$\mathbb{E} \left[\int_0^T \|\zeta(t, X_t^\zeta) - \hat{\zeta}_h(t, X_t^\zeta)\|_{\mathbb{R}^\ell}^2 dt \right] \leq Ch.$$