A PROOFS

A.1 Deterministic Setting

The following proof follows nearly the same proof as the main result in Argyros (1999) with a few minor modifications in the conclusion; we provide it here for posterity.

Proof [Proof Proposition 3] Since $\|I - \Gamma D\omega(x)\| < 1$ for each $x \in B_{r_0}(x^*)$, as stated in the proposition statement, there exists $0 < r' < r'' < 1$ such that $\|I - \Gamma D\omega(x)\| \leq r' < r'' < 1$ for all $x \in B_r(x^*)$. Since

$$\lim_{x \rightarrow x^*} \|R(x - x^*)\|/\|x - x^*\| = 0,$$

for $0 < 1 - r'' < 1$, there exists $\bar{r} > 0$ such that

$$\|R(x - x^*)\| \leq (1 - r'')\|x - x^*\|, \forall x \in B_{\bar{r}}(x^*).$$

As in the proposition statement, let $r$ be the largest, finite such $\bar{r}$. Note that for arbitrary $c > 0$, there exists $\tilde{r} > 0$ such that the bound on $\|R(x - x^*)\|$ holds; hence, we choose $c = 1 - r''$ and find the largest such $\tilde{r}$ for which the bound holds. Combining the above bounds with the definition of $g$, we have that

$$\|g(x) - g(x^*)\| \leq (1 - \delta)\|x - x^*\|, \forall x \in B_\tilde{r}(x^*)$$

where $\delta = r'' - r'$ and $r^* = \min\{r_0, r\}$. Hence, applying the result iteratively, we have that

$$\|x_t - x^*\| \leq (1 - \delta)^t\|x_0 - x^*\|, \forall x_0 \in B_{\tilde{r}}(x^*).$$

Note that $0 < 1 - \delta < 1$. Using the approximation $1 - \delta < \exp(-\delta)$, we have that

$$\|x_T - x^*\| \leq \exp(-T\delta)\|x_0 - x^*\|$$

so that $x_t \in B_c(x^*)$ for all $t \geq T = \lceil \delta^{-1} \log(r^*/\epsilon) \rceil$.

As noted in the remark, a similar result holds under the relaxed assumption that $\rho(I - \Gamma D\omega(x)) < 1$ for all $x \in B_{r_0}(x^*)$. To see this, we first note that $\rho(I - \Gamma D\omega(x)) < 1$ implies there exists $c > 0$ such that $\rho(I - \Gamma D\omega(x)) \leq c < 1$. Hence, given any $\epsilon > 0$, there is a norm on $\mathbb{R}^d$ and a $c > 0$ such that $\|I - \Gamma D\omega\| \leq c + \epsilon < 1$ on $B_{r_0}(x^*)$ (Ortega and Rheinboldt, 1970, 2.2.8). Then, we can apply the same argument as above using $r' = c + \epsilon$.

A.2 Stochastic Setting

A key tool used in the finite-time two-timescale analysis is the nonlinear variation of constants formula of Alekseev Alekseev (1961), Borkar and Pattathil (2018).

Theorem 1. Consider a differential equation

$$\dot{u}(t) = f(t, u(t)), \; t \geq 0,$$

and its perturbation

$$\dot{p}(t) = f(t, p(t)) + g(t, p(t)), \; t \geq 0$$

where $f, g : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $f \in C^1$, and $g \in C$. Let $u(t, t_0, p_0)$ and $p(t, t_0, p_0)$ denote the solutions of the above nonlinear systems for $t \geq t_0$ satisfying $u(t_0, t_0, p_0) = p(t_0, t_0, p_0) = p_0$, respectively. Then,

$$p(t, t_0, p_0) = u(t, t_0, p_0) + \int_{t_0}^{t} \Phi(t, s, p(s, t_0, p_0)) \cdot g(s, p(s, t_0, p_0)) ds, \; t \geq t_0$$

where $\Phi(t, s, u_0)$, for $u_0 \in \mathbb{R}^d$, is the fundamental matrix of the linear system

$$\dot{v}(t) = \frac{\partial f}{\partial u}(t, u(t, s, u_0))v(t), \; t \geq s \tag{1}$$

with $\Phi(s, s, u_0) = I_d$, the $d$-dimensional identity matrix.

Consider a locally asymptotically stable differential Nash equilibrium $x^* = (\lambda(x^*_2), x^*_1) \in X$ and let $B_{r_0}(x^*)$ be an $r_0 > 0$ radius ball around $x^*$ contained in the region of attraction. Stability implies that the Jacobian $J_S(\lambda(x^*_2), x^*_1)$ is positive definite and by the converse Lyapunov theorem (Sastry, 1999, Chap. 5) there exists local Lyapunov functions for the dynamics $\tilde{x}_2(t) = -\tau D_2f_2(\lambda(x_2(t)), x_2(t))$ and for the dynamics $\tilde{x}_1(t) = -D_1f_1(x_1(t), x_2(t))$, for each fixed $x_2$. In particular, there exists a local Lyapunov function $V \in C^1(\mathbb{R}^d)$ with $\lim_{\|x_2\| \rightarrow \infty} V(x_2) = \infty$, and $(\nabla V(x_2), D_2f_2(\lambda(x_2), x_2)) < 0$ for $x_2 \neq x^*_2$. For $r > 0$, let $V^r = \{x \in \text{dom}(V) : V(x) \leq r\}$. Then, there is also $r > r_0 > 0$ and $\epsilon_0 > 0$ such that for each $r < \epsilon_0$, $\{x_2 \in \mathbb{R}^d \mid \|x_2 - x_2^*\| \leq \epsilon \} \subseteq V^{r_0} \subset N_{\epsilon_0}(V^{r_0}) \subseteq V^r \subset \text{dom}(V)$ where $N_{\epsilon_0}(V^{r_0}) = \{x \in \mathbb{R}^d \mid \exists x' \in V^{r_0} \text{s.t.} \|x' - x\| \leq \epsilon_0\}$. An analogously defined $\tilde{V}$ exists for the dynamics $\tilde{x}_1$ for each fixed $x_2$.

For now, fix $n_0$ sufficiently large; we specify this a bit further down. Define the event $\mathcal{E}_{\tilde{n}} = \{\tilde{x}_1(t) \in V^r \forall t \in [\tilde{t}_{n_0}, \tilde{t}_0] \}$ where $\tilde{x}_1(t) = x_1(t) + \frac{\tilde{t}_0 - \tilde{t}_1}{\tilde{t}_1}(x_1(t_1) - x_1(t))$ are linear interpolates defined for $t \in (\tilde{t}_k, \tilde{t}_{k+1})$ with $\tilde{t}_{k+1} = \tilde{t}_k + \gamma_{1,k}$ and $\tilde{t}_0 = 0$. The basic idea of the proof is to leverage Alekseev’s formula (Theorem 1) to bound the difference between the linearly interpolated trajectories (i.e., asymptotic pseudo-trajectories) and the flow of the corresponding limiting differential equation on each continuous time interval between each of the successive iterates $k$ and $k + 1$ by a number that decays asymptotically. Then, for large enough $n$, a union bound is used.
over all the remaining time intervals to construct a concentration bound. This is done first for fast player (i.e. player 1), to show that $x_{1,k}$ tracks $\lambda(x_{2,k})$, and then for the slow player (i.e. player 2).

Following Borkar and Pattathil (2018), we can express the linear interpolates for any $n \geq n_0$ as $\tilde{x}_1(t_{n+1}) = \tilde{x}_1(t_n) - \gamma_{1,k} D_1f_1(x_k) + w_{1,k+1}$, where

$$
\gamma_{1,k} D_1 f_1(x_k) = \int_{t_k}^{t_{k+1}} D_1 f_1(\tilde{x}_1(t), x_2) dt
$$

and similarly for the $w_{1,k+1}$ term. Adding and subtracting $\int_{t_n}^{t_{n+1}} D_1 f_1(\tilde{x}_1(s), x_2(s))$, Alekseev’s formula can be applied to get

$$
\tilde{x}_1(t) = x_1(t) + \Phi_1(t, s, \tilde{x}_1(t_n), x_2(t_n))(\tilde{x}_1(t_n)) - x_1(t_n) + \int_s^t \Phi_2(t, s, \tilde{x}_1(s), x_2(s)) \zeta_1(s) ds
$$

where $x_2(t) \equiv x_2$ is constant (since $\dot{x}_2 = 0$), $x_1(t) = \lambda(x_2)$, $\zeta_1(s) = -D_1 f_1(\tilde{x}_1(t_k), x_2(t_k)) + D_1 f_1(\tilde{x}_1(s), x_2(s)) + w_{1,k+1}$, and where for $t \geq s$, $\Phi_1(\cdot)$ satisfies linear system

$$
\dot{\Phi}_1(t, s, x_0) = J_1(x_1(t), x_2(t)) \Phi_1(t, s, x_0),
$$

with initial data $\Phi_1(t, s, x_0) = I$ and $x_0 = (x_{1,0}, x_{2,0})$ and where $J_1$ the Jacobian of $-D_1 f_1(\cdot, x_2)$.

Given that $x^* = (\lambda(x_2^*), x_2^*)$ is a stable differential Nash equilibrium, $J_1(x^*)$ is positive definite. Hence, as in (Thoppe and Borkar, 2018, Lem. 5.3), we can find $M_0, \kappa_1 > 0$ such that for $t \geq s$, $x_{1,0} \in V$, $\|\Phi_1(t, s, x_{1,0}, x_{2,0})\| \leq M e^{-\kappa_1(t-s)}$; this result follows from standard results on stability of linear systems (see, e.g., Callier and Desoer (1991, §7.2, Thm. 33)) along with a bound on $\int_s^t \|D^2 f_1(x_1, x_2(\tau, \tilde{x})) - D^2 f_1(x^*)\| d\tau$ for $\tilde{x} \in V^r$ (see, e.g., (Thoppe and Borkar, 2018, Lem 5.2)).

Consider $z_k = \lambda(x_{2,k})$—i.e., where $D_1 f_1(x_{1,k}, x_{2,k}) = 0$. Then, using a Taylor expansion of the implicitly defined $\lambda$, we get

$$
z_{k+1} = z_k + D\lambda(x_{2,k})(x_{2,k+1} - x_{2,k}) + \delta_{k+1}
$$

where $\|\delta_{k+1}\| \leq L_e \|x_{2,k+1} - x_{2,k}\|^2$ is the error from the remainder terms. Plugging in $x_{2,k+1},$

$$
z_{k+1} = z_k + \gamma_{1,k}(-D_1 f_1(x_{1,k}, x_{2,k}) + \tau_k \lambda(x_{2,k}) + (w_{2,k+1} - D_2 f_2(x_{1,k}, x_{2,k})) + \gamma_{1,k}^{-1} \delta_{k+1})
$$

The terms after $-D_1 f_1$ are $o(1)$, and hence asymptotically negligible, so that this $z$ sequence tracks dynamics as $x_{1,k}$. We show that with high probability, they asymptotically contract to one another.

Now, let us bound the normed difference between $x_{1,k}$ and $z_k$. Define constant $H_{n_0} = (\|\tilde{x}_1(t_n) - x_1(t_n)\| + \|\tilde{z}(t_n) - x_1(t_n)\|)$ and

$$
S_{1,n} = \sum_{k=n_0}^{n-1} \int_{t_k}^{t_{k+1}} \Phi_1(t_n, s, \tilde{x}_1(t_k), x_2(t_k)) ds
$$

Let $\tau_k = \gamma_{2,k}/\gamma_{1,k}$.

**Lemma 1.** For any $n \geq n_0$, there exists $K > 0$ such that

$$
\|x_{1,n} - z_n\| \leq K \left(\|S_{1, n}\| + e^{-\kappa_1(t_n - t_{n_0})} H_{n_0}ight) + sup_{n_0 \leq k \leq n-1} \gamma_{1,k} + sup_{n_0 \leq k \leq n-1} \gamma_{1,k} \|w_{1,k+1}\|^2
$$

conditioned on $\mathcal{E}_n$.

In order to construct a high-probability bound for $x_{2,k}$, we need a similar bound as in Lemma 1 can be constructed for $x_{2,k}$.

Define the event $\mathcal{E}_n = \{ x_2(t) \in V^r \ \forall t \in [t_n, t_{n+1}] \}$ where $x_2(t) = x_{2,k} + \frac{1}{\gamma_{2,k}} (x_{2,k+1} - x_{2,k})$ is the linear interpolated points between the samples $\{x_{2,k}\}, \tilde{t}_k + 1 = \tilde{t}_k + \gamma_{1,k}$, and $\tilde{t}_0 = 0$. Then as above, Alekseev’s formula can again be applied to get

$$
\tilde{x}_2(t) = x_2(t, \tilde{t}_n, x_2(t_{n_0})) + \Phi_2(t, \tilde{t}_n, \tilde{x}_2(t_{n_0}))
$$

where $x_2(t) \equiv x_2^*$,

$$
\zeta_2(s) = D_2 f_2(\lambda(x_{2,k}), x_{2,k}) - D_2 f_2(\lambda(x_2^*), x_2^*) + D_2 f_2(x_k) - D_2 f_2(\lambda(x_{2,k}), x_{2,k}) + w_{2,k+1},
$$

and $\Phi_2$ is the solution to a linear system with dynamics $J_2(\lambda(x_2^*), x_2^*)$, the Jacobian of $-D_2 f_2(\lambda(\cdot), \cdot)$, and with initial data $\Phi_2(s, x_{2,0}) = I$. This linear system, as above, has bound $\|\Phi_2(s, x_{2,0})\| \leq M_2 e^{\kappa_2(t_1)}$ for some $M_2, \kappa_2 > 0$. Define

$$
S_{2,n} = \sum_{k=n_0}^{n-1} \int_{t_k}^{t_{k+1}} \Phi_2(t_n, s, \tilde{x}_2(t_k)) ds
$$

conditioned on $\mathcal{E}_n$.

Using this lemma, we can get the desired guarantees on $x_{1,k}$. 
A.3 Uniform Learning Rates

Before concluding, we specialize to the case in which agents have the same learning rate sequence $\gamma_{i,k} = \gamma_k$ for each $i \in I$.

**Theorem 2.** Suppose that $x^*$ is a stable differential Nash equilibrium of the game $(f_1, \ldots, f_n)$, and that Assumption 2 holds (excluding A2b.iii). For each $n_i$, let $n_0 \geq 0$ and $\zeta_n = \max_{n_0 \leq k \leq n-1} \left( \text{exp}(\lambda \sum_{i=0}^{k-1} \gamma_i) \right)$. Given any $\epsilon > 0$, such that $B_r(x^*) \subset B_r(x^*) \subset \mathcal{V}$, there exists constants $C_1, C_2 > 0$ and functions $h_1(\epsilon) = O(\log(1/\epsilon))$ and $h_2(\epsilon) = O(1/\epsilon)$ so that whenever $T \geq h_1(\epsilon)$ and $n_0 \geq N$, where $N$ is such that $1/\gamma_n \geq h_2(\epsilon)$ for all $n \geq N$, the samples generated by the gradient-based learning rule satisfy

$$
\Pr(\bar{x}(t) \in B_\epsilon(x^*) \forall t \geq t_n + T + 1 | \bar{x}(n_0) \in B_r(x^*)) \geq 1 - \sum_{n=n_0}^{\infty} (C_1 \exp(-C_2 \epsilon^{1/2}/\gamma_n^{1/2}) + C_1 \exp(-C_2 \min\{\epsilon, \epsilon^2\}/\zeta_n))
$$

where the constants depend only on parameters $\lambda, r, \tau_L$ and the dimension $d = \sum_i d_i$. Then stochastic gradient-based learning in games attains an $\epsilon$-stable differential Nash $x^*$ in finite time with high probability.

The above theorem implies that $x_k \in B_\epsilon(x^*)$ for all $k \geq n_0 + \lceil \log(4K/\epsilon) \lambda^{-1} \rceil + 1$ with high probability for some constant $K$ that depends only on $\lambda, r, \tau_L$, and $d$.

**Proof** Since $x^*$ is a stable differential Nash equilibrium, $D_i^2 f_i(x^*)$ is positive definite and $D_i^2 f_i(x^*)$ is positive definite for each $i \in I$. Thus $x^*$ is a locally asymptotically stable hyperbolic equilibrium point of $\dot{x} = -\omega(x)$. Hence, the assumptions of Theorem 1.1 Thoppe and Borkar (2018) are satisfied so that we can invoke the result which gives us the high probability bound for stochastic gradient-based learning in games. □

B ADDITIONAL EXAMPLES

In this section, we provide additional numerical examples.

B.1 Matching pennies

This example is a classic bimatrix game—matching pennies—where agents have zero-sum costs associated with the matrices $(A, B)$ below. We parameterize agents with a “soft” arg max policy where they play smoothed best-response. This game has been well studied in the game theory literature and we use this example illustrate the warping of agent’s vector field under non-uniform learning rates.

Consider the zero-sum bimatrix game with

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

where each agent’s costs are $f_1(x,y) = \pi(y)^T A \pi(x)$ and $f_2(x,y) = \pi(x)^T B \pi(y)$, and soft max policy as

$$\pi(z) = \frac{e^{10z}}{e^{10z} + e^{10(1-z)}}, \quad \frac{e^{10(1-z)}}{e^{10z} + e^{10(1-z)}} \right).$$

The mixed Nash equilibrium for this game is $(x^*, y^*) = (0.5, 0.5)$, but the Jacobian of the gradient dynamics at this fixed point is

$$J(x^*, y^*) = \begin{bmatrix} 0 & 100 \\ -100 & 0 \end{bmatrix}$$

and has purely imaginary eigenvalues $\pm 100i$, therefore admits a limit cycle. Regardless, we can visualize the effects of non-uniform learning rates to the gradient dynamics in Figure 1. We notice that the gradient flow stretches along the axes of the faster agent (the agent with a larger learning rate). However, the fixed point of these dynamics does not change.

References


