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

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

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

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

As in the proposition statement, let r be the largest, finite such $\tilde{r}$. 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

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

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

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

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

■

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 $r'=c+\varepsilon$.

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*

$$\dot{u}(t)=f(t,u(t)),\ t\geq0,$$

and its perturbation

$$\dot{p}(t)=f(t,p(t))+g(t,p(t)),\ t\geq0$$

where $f,g:\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}^d$, $f\in C^1$, 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 $x^*=(\lambda(x_2^*),x_2^*)\in X$ 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 $J_{\mathcal{S}}(\lambda(x_2^*),x_2^*)$ is positive definite and by the converse Lyapunov theorem (Sastry, 1999, Chap. 5) there exists local Lyapunov functions for the dynamics $\dot{x}_2(t)=-\tau D_2f_2(\lambda(x_2(t)),x_2(t))$ and for the dynamics $\dot{x}_1(t)=-D_1f_1(x_1(t),x_2)$, for each fixed x_2. In particular, there exists a local Lyapunov function $V\in C^1(\mathbb{R}^{d_1})$ with $\lim_{\|x_2\|\uparrow\infty}V(x_2)=\infty$, and $\langle\nabla V(x_2),D_2f_2(\lambda(x_2),x_2)\rangle<0$ for $x_2\neq x_2^*$. For $r>0$, let $V^r=\{x\in\mathrm{dom}(V):\ V(x)\leq r\}$. Then, there is also $r>r_0>0$ and $\epsilon_0>0$ such that for $\epsilon<\epsilon_0$, $\{x_2\in\mathbb{R}^{d_2}|\ \|x_2-x_2^*\|\leq\epsilon\}\subseteq V^{r_0}\subset\mathcal{N}_{\epsilon_0}(V^{r_0})\subseteq V^r\subset\mathrm{dom}(V)$ where $\mathcal{N}_{\epsilon_0}(V^{r_0})=\{x\in\mathbb{R}^{d_2}|\ \exists x'\in V^{r_0}\text{ s.t.}\|x'-x\|\leq\epsilon_0\}$. An analogously defined $\tilde{V}$ exists for the dynamics $\dot{x}_1$ for each fixed x_2.

For now, fix n_0 sufficiently large; we specify this a bit further down. Define the event $\mathcal{E}_n=\{\bar{x}_1(t)\in V^r\ \forall t\in[\tilde{t}_{n_0},\tilde{t}_n]\}$ where $\bar{x}_1(t)=x_{1,k}+\frac{t-\tilde{t}_k}{\gamma_{1,k}}(x_{1,k+1}-x_{1,k})$ 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 psuedo-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 $\bar{x}_1(\tilde{t}_{n+1}) = \bar{x}_1(\tilde{t}_{n_0}) - \sum_{k=n_0}^{n} \gamma_{1,k}(D_1 f_1(x_k) + w_{1,k+1})$ where

$$\gamma_{1,k} D_1 f_1(x_k) = \int_{\tilde{t}_k}^{\tilde{t}_{k+1}} D_1 f_1(\bar{x}_1(\tilde{t}_k), x_{2,k})$$

and similarly for the $w_{1,k+1}$ term. Adding and subtracting $\int_{\tilde{t}_{n_0}}^{\tilde{t}_{n+1}} D_1 f_1(\bar{x}_1(s), x_2(s),)$, Alekseev's formula can be applied to get

$$\begin{aligned}\bar{x}_1(t) = x_1(t) + \Phi_1(t, s, \bar{x}_1(\tilde{t}_{n_0}), x_2(\tilde{t}_{n_0}))(\bar{x}_1(\tilde{t}_{n_0}) \\ - x_1(\tilde{t}_{n_0})) + \int_{\tilde{t}_{n_0}}^{t} \Phi_2(t, s, \bar{x}_1(s), x_2(s))\zeta_1(s)\, ds\end{aligned}$$

where $x_2(t) \equiv x_2$ is constant (since $\dot{x}_2 = 0$), $x_1(t) = \lambda(x_2)$, $\zeta_1(s) = -D_1 f_1(\bar{x}_1(\tilde{t}_k), x_2(\tilde{t}_k)) + D_1 f_1(\bar{x}_1(s), x_2(s)) + w_{1,k+1}$, 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, $\kappa_1 > 0$ such that for $t \geq s$, $x_{1,0} \in V^r$, $\|\Phi_1(t, s, x_{1,0}, x_{2,0})\| \leq M e^{-\kappa_1(t-s)}$; 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 $\int_s^t \|D_1^2 f_1(x_1, x_2(\tau, s, \tilde{x}_0)) - D_1^2 f_1(x^*)\| d\tau$ for $\tilde{x}_0 \in V^r$ (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 λ, we get

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

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

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

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 $H_{n_0} = (\|\bar{x}_1(\tilde{t}_{n_0} - x_1(\tilde{t}_{n_0})\| + \|\bar{z}(\tilde{t}_{n_0}) - x_1(\tilde{t}_{n_0})\|)$ and

$$\begin{aligned}S_{1,n} = \textstyle\sum_{k=n_0}^{n-1} \big(\int_{\tilde{t}_k}^{\tilde{t}_{k+1}} \Phi_1(\tilde{t}_n, s, \bar{x}_1(\tilde{t}_k), x_2(\tilde{t}_k)) ds\big) \\ \cdot w_{2,k+1}.\end{aligned}$$

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

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

$$\begin{aligned}&\|x_{1,n} - z_n\| \leq K\big(\|S_{1,n}\| + e^{-\kappa_1(\tilde{t}_n - \tilde{t}_{n_0})} H_{n_0} \\ &+ \sup_{n_0 \leq k \leq n-1} \gamma_{1,k} + \sup_{n_0 \leq k \leq n-1} \gamma_{1,k} \|w_{1,k+1}\|^2 \\ &+ \sup_{n_0 \leq k \leq n-1} \tau_k + \sup_{n_0 \leq k \leq n-1} \tau_k \|w_{2,k+1}\|^2\big)\end{aligned}$$

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 $\hat{\mathcal{E}}_n = \{\bar{x}_2(t) \in V^r\ \forall t \in [\hat{t}_{n_0}, \hat{t}_n]\}$ where $\bar{x}_2(t) = x_{2,k} + \frac{t - \hat{t}_k}{\gamma_{2,k}}(x_{2,k+1} - x_{2,k})$ is the linear interpolated points between the samples $\{x_{2,k}\}$, $\hat{t}_{k+1} = \hat{t}_k + \gamma_{1,k}$, and $\hat{t}_0 = 0$. Then as above, Alekseev's formula can again be applied to get

$$\begin{aligned}&\bar{x}_2(t) = x_2(t, \hat{t}_{n_0}, x_2(\hat{t}_{n_0})) + \Phi_2(t, \hat{t}_{n_0}, \bar{x}_2(\hat{t}_{n_0})) \\ &\quad \cdot (\bar{x}_2(\hat{t}_{n_0}) - x_2(\hat{t}_{n_0})) + \int_{\hat{t}_{n_0}}^{t} \Phi_2(t, s, \bar{x}_2(s))\zeta_1(s)\, ds\end{aligned}$$

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

$$\begin{aligned}\zeta_1(s) = D_2 f_2(\lambda(x_{2,k}), x_{2,k}) - D_2 f_2(\lambda(\bar{x}_2(s)), \bar{x}_2(s)) \\ + D_2 f_2(x_k) - D_2 f_2(\lambda(x_{2,k}), x_{2,k}) + w_{2,k+1},\end{aligned}$$

and Φ_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 $\Phi_2(s, s, x_{2,0}) = I$. This linear system, as above, has bound $\|\Phi_2(t, s, x_{2,0})\| \leq M_2 e^{\kappa_2(t-1)}$ for some $M_2, \kappa_2 > 0$. Define

$$\begin{aligned}S_{2,n} = \textstyle\sum_{k=n_0}^{n-1} \int_{\hat{t}_k}^{\hat{t}_{k+1}} \Phi_2(\hat{t}_n, s, \bar{x}_2(\hat{t}_k)) ds \\ \cdot w_{2,k+1}.\end{aligned}$$

Lemma 2. *For any $n \geq n_0$, there exists $\bar{K} > 0$ such that*

$$\begin{aligned}&\|\bar{x}_2(\hat{t}_n) - x_2(\hat{t}_n)\| \leq \bar{K}\big(\|S_{2,n}\| + \sup_{n_0 \leq k \leq n-1} \|S_{1,k}\| \\ &+ \sup_{n_0 \leq k \leq n-1} \gamma_{1,k} + \sup_{n_0 \leq k \leq n-1} \gamma_{1,k} \|w_{1,k+1}\|^2 \\ &+ \sup_{n_0 \leq k \leq n-1} \tau_k + \sup_{n_0 \leq k \leq n-1} \tau_k \|w_{2,k+1}\|^2 \\ &+ e^{\kappa_2(\hat{t}_n - \hat{t}_{n_0})} \|\bar{x}_2(\hat{t}_{n_0}) - x_2(\hat{t}_{n_0})\| \\ &+ \sup_{n_0 \leq k \leq n-1} \tau_k H_{n_0}\big)\end{aligned}$$

conditioned on $\tilde{\mathcal{E}}_n$.

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 $i \in \mathcal{I}$.

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, let $n_0 \geq 0$ and $\zeta_n = \max_{n_0 \leq k \leq n-1} \big(\exp(-\lambda \sum_{\ell=k+1}^{n-1} \gamma_\ell\big)\gamma_k$. Given any $\varepsilon > 0$ such that $B_\varepsilon(x^*) \subset B_r(x^*) \subset \mathcal{V}$, there exists constants $C_1, C_2 > 0$ and functions $h_1(\varepsilon) = O(\log(1/\varepsilon))$ and $h_2(\varepsilon) = O(1/\varepsilon)$ so that whenever $T \geq h_1(\varepsilon)$ and $n_0 \geq N$, where N is such that $1/\gamma_n \geq h_2(\varepsilon)$ for all $n \geq N$, the samples generated by the gradient-based learning rule satisfy*

$$\begin{aligned}
&\Pr\left(\bar{x}(t) \in B_\varepsilon(x^*)\ \forall t \geq t_{n_0} + T + 1 |\ \bar{x}(t_{n_0}) \in B_r(x^*)\right) \\
&\geq 1 - \textstyle\sum_{n=n_0}^{\infty} \big(C_1 \exp(-C_2 \varepsilon^{1/2}/\gamma_n^{1/2}) \\
&\quad + C_1 \exp(-C_2 \min\{\varepsilon, \varepsilon^2\}/\zeta_n)\big)
\end{aligned}$$

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

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

Proof Since x^* is a stable differential Nash equilibrium, $J(x^*)$ is positive definite and $D_i^2 f_i(x^*)$ is positive definite for each $i \in \mathcal{I}$. 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.

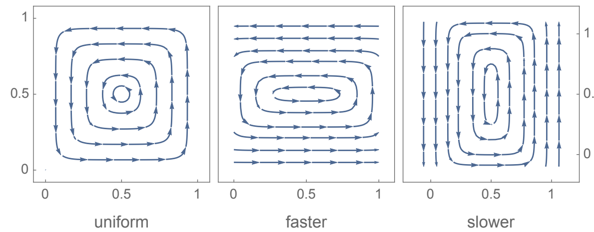

Figure 1: Gradient dynamics of the matching pennies game where agents learning have different learning rates. The vector field of the gradient dynamics are stretched along the faster agent's coordinate.

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

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

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

V. M. Alekseev. An estimate for the perturbations of the solutions of ordinary differential equations. *Vestnik Moskov. Univ. Ser. I. Mat. Meh.*, 2:28–36, 1961.

I. K. Argyros. A generalization of ostrowski's theorem on fixed points. *Applied Mathematics Letters*, 12:77–79, 1999.

Vivek S. Borkar and Sarath Pattathil. Concentration bounds for two time scale stochastic approximation. *arxiv:1806.10798*, 2018.

F. Callier and C. Desoer. *Linear Systems Theory*. Springer, 1991.

J. M. Ortega and W. C. Rheinboldt. *Iterative Solutions to Nonlinear Equations in Several Variables*. Academic Press, 1970.

Shankar Sastry. *Nonlinear Systems*. Springer New York, 1999. doi: 10.1007/978-1-4757-3108-8.

G. Thoppe and V. S. Borkar. A concentration bound for stochastic approximation via alekseev's formula. *arXiv:1506.08657v3*, 2018.