

## A Omitted Proofs

In this section we include all the proofs deferred from the main body. For the convenience of the reader, all claims will be restated. Before we proceed with the proof of Theorem 2.3, let us point out some additional notational conventions. First, with a standard abuse of notation, we overload

$$u_i : \times_{i=1}^p \Delta(\mathcal{A}_i) \ni (\mathbf{x}_1, \dots, \mathbf{x}_p) \mapsto \mathbb{E}_{\mathbf{a} \sim \mathbf{x}} [u_i(\mathbf{a})] = \sum_{\mathbf{a} \in \mathcal{A}} u_i(a_1, \dots, a_p) \prod_{j \in [p]} \mathbf{x}_j(a_j)$$

to denote the mixed extension of player's  $i \in [p]$  utility function, where  $\mathcal{A} := \times_{j=1}^p \mathcal{A}_j$ . Furthermore, for  $a_i \in \mathcal{A}_i$  we let

$$u_i(a_i, \mathbf{x}_{-i}) := \sum_{\mathbf{a}_{-i} \in \mathcal{A}_{-i}} u_i(a_1, \dots, a_i, \dots, a_p) \prod_{j \neq i} \mathbf{x}_j(a_j),$$

where  $\mathcal{A}_{-i} := \times_{j \neq i} \mathcal{A}_j$ .

**Theorem 2.3** (Folklore). *Suppose that every player  $i \in [p]$  employs a no-regret learning algorithm with regret  $\text{Reg}_i^T$  up to time  $T \in \mathbb{N}$ . Moreover, let  $\boldsymbol{\mu}^{(t)} := \mathbf{x}_1^{(t)} \otimes \dots \otimes \mathbf{x}_p^{(t)}$  be the correlated distribution of play at time  $t \in [T]$ , and  $\bar{\boldsymbol{\mu}} := \frac{1}{T} \sum_{t=1}^T \boldsymbol{\mu}^{(t)}$  be the average correlated distribution of play up to time  $T$ . Then,*

$$\mathbb{E}_{\mathbf{a} \sim \bar{\boldsymbol{\mu}}} [u_i(a'_i, \mathbf{a}_{-i})] \leq \mathbb{E}_{\mathbf{a} \sim \bar{\boldsymbol{\mu}}} [u_i(\mathbf{a})] + \frac{1}{T} \max_{i \in [p]} \text{Reg}_i^T.$$

*Proof.* By definition of regret (1), it follows that for any player  $i \in [p]$  and any possible deviation  $a'_i \in \mathcal{A}_i$ ,

$$\sum_{t=1}^T u_i(a'_i, \mathbf{x}_{-i}^{(t)}) - \sum_{t=1}^T u_i(\mathbf{x}^{(t)}) \leq \text{Reg}_i^T. \quad (6)$$

Moreover,

$$\mathbb{E}_{\mathbf{a} \sim \bar{\boldsymbol{\mu}}} [u_i(\mathbf{a})] = \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{\mathbf{a} \sim \boldsymbol{\mu}^{(t)}} [u_i(\mathbf{a})] = \frac{1}{T} \sum_{t=1}^T u_i(\mathbf{x}^{(t)});$$

and

$$\mathbb{E}_{\mathbf{a} \sim \bar{\boldsymbol{\mu}}} [u_i(a'_i, \mathbf{a}_{-i})] = \frac{1}{T} \sum_{t=1}^T \mathbb{E}_{\mathbf{a} \sim \boldsymbol{\mu}^{(t)}} [u_i(a'_i, \mathbf{a}_{-i})] = \frac{1}{T} \sum_{t=1}^T u_i(a'_i, \mathbf{x}_{-i}^{(t)}).$$

Thus, the theorem follows directly from (6).  $\square$

**Corollary 3.1.** *Suppose that both players employ (OMD) with learning rate  $\eta > 0$ . Then,*

$$\begin{aligned} \text{Reg}_{\mathcal{X}}^T &\leq \frac{\Omega_{\mathcal{R}\mathcal{X}}}{\eta} + \eta \|\mathbf{A}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 - \frac{1}{4\eta} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t-1)}\|^2 \right); \\ \text{Reg}_{\mathcal{Y}}^T &\leq \frac{\Omega_{\mathcal{R}\mathcal{Y}}}{\eta} + \eta \|\mathbf{B}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 - \frac{1}{4\eta} \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|^2 \right). \end{aligned}$$

*Proof.* First, the utility  $\mathbf{u}_{\mathcal{X}}^{(t)}$  observed by player  $\mathcal{X}$  at time  $t \geq 0$  is equal to  $\mathbf{A}\mathbf{y}^{(t)}$ . Thus, the claimed bound for  $\text{Reg}_{\mathcal{X}}^T$  follows directly from Proposition 2.1 using the fact that  $\|\mathbf{A}\mathbf{y}^{(t)} - \mathbf{A}\mathbf{y}^{(t-1)}\|_* \leq \|\mathbf{A}\|_{\text{op}} \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|$ . Similarly, the utility  $\mathbf{u}_{\mathcal{Y}}^{(t)}$  observed by player  $\mathcal{Y}$  at time  $t \geq 0$  is tantamount to  $\mathbf{B}^\top \mathbf{x}^{(t)}$ . As a result, the bound for  $\text{Reg}_{\mathcal{Y}}^T$  follows from Proposition 2.1 and the fact that  $\|\mathbf{B}^\top \mathbf{x}^{(t)} - \mathbf{B}^\top \mathbf{x}^{(t-1)}\|_* \leq \|\mathbf{B}\|_{\text{op}} \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|$ , since  $\|\mathbf{B}^\top\|_{\text{op}} = \|\mathbf{B}\|_{\text{op}}$ .  $\square$

**Proposition 3.2** (Approximate Fixed Points of OMD). *Consider a bimatrix game  $(\mathbf{A}, \mathbf{B})$ , and suppose that both players employ (OMD) with learning rate  $\eta > 0$  and a  $G$ -smooth regularizer. Then, if  $\|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t-1)}\|, \|\hat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\| \leq \epsilon\eta$  and  $\|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|, \|\hat{\mathbf{y}}^{(t)} - \mathbf{y}^{(t)}\| \leq \epsilon\eta$ , the pair  $(\mathbf{x}^{(t)}, \mathbf{y}^{(t)})$  is a  $(2\epsilon G \max\{\Omega_{\mathcal{X}}, \Omega_{\mathcal{Y}}\} + \epsilon\eta)$ -approximate Nash equilibrium of  $(\mathbf{A}, \mathbf{B})$ .*

*Proof.* First, by definition of the Bregman divergence, the update rule of (OMD) can be equivalently expressed as

$$\begin{aligned}\mathbf{x}^{(t)} &:= \arg \max_{\mathbf{x} \in \mathcal{X}} \left\{ \langle \mathbf{x}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle - \frac{1}{\eta} \mathcal{R}(\mathbf{x}) + \frac{1}{\eta} \langle \mathbf{x}, \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t-1)}) \rangle \right\}; \\ \widehat{\mathbf{x}}^{(t)} &:= \arg \max_{\widehat{\mathbf{x}} \in \mathcal{X}} \left\{ \langle \widehat{\mathbf{x}}, \mathbf{A}\mathbf{y}^{(t)} \rangle - \frac{1}{\eta} \mathcal{R}(\widehat{\mathbf{x}}) + \frac{1}{\eta} \langle \widehat{\mathbf{x}}, \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t-1)}) \rangle \right\}.\end{aligned}$$

Now the maximization problem associated with the update rule of the secondary sequence can be equivalently cast in a variational inequality form (e.g., see [Facchinei and Pang, 2003]):

$$\left\langle \widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} - \frac{1}{\eta} \left( \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t)}) - \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t-1)}) \right) \right\rangle \leq 0, \quad \forall \widehat{\mathbf{x}} \in \mathcal{X}.$$

Thus,

$$\begin{aligned}\langle \widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle &\leq \frac{1}{\eta} \langle \widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}, \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t)}) - \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t-1)}) \rangle \\ &\leq \frac{1}{\eta} \|\widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}\| \|\nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t)}) - \nabla \mathcal{R}(\widehat{\mathbf{x}}^{(t-1)})\|_* \quad (7)\end{aligned}$$

$$\leq \frac{G}{\eta} \|\widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}\| \|\widehat{\mathbf{x}}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \quad (8)$$

$$\leq 2\epsilon G \Omega_{\mathcal{X}}, \quad (9)$$

for any  $\widehat{\mathbf{x}} \in \mathcal{X}$ , where (7) derives from the Cauchy-Schwarz inequality; (8) follows since  $\mathcal{R}_{\mathcal{X}}$  is assumed to be  $G$ -smooth; and (9) uses that  $\|\widehat{\mathbf{x}}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq 2\epsilon\eta$ , which in turn follows since  $\|\widehat{\mathbf{x}}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\| + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq 2\epsilon\eta$  (triangle inequality), as well as the fact that, by definition,  $\|\widehat{\mathbf{x}} - \widehat{\mathbf{x}}^{(t)}\| \leq \Omega_{\mathcal{X}}$  for any  $\widehat{\mathbf{x}} \in \mathcal{X}$ . As a result, we have shown that for any  $\widehat{\mathbf{x}} \in \mathcal{X}$ ,

$$\langle \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle \geq \langle \widehat{\mathbf{x}}, \mathbf{A}\mathbf{y}^{(t)} \rangle - 2\epsilon G \Omega_{\mathcal{X}}. \quad (10)$$

Furthermore, by Cauchy-Schwarz inequality we have that

$$\langle \mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle \geq -\|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\| \|\mathbf{A}\mathbf{y}^{(t)}\|_* \geq -\epsilon\eta, \quad (11)$$

where we used the normalization assumption  $\|\mathbf{A}\mathbf{y}^{(t)}\|_* \leq 1$ . Thus, combing (11) with (10) yields that

$$\langle \mathbf{x}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle \geq \langle \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle - \epsilon\eta \geq \langle \widehat{\mathbf{x}}, \mathbf{A}\mathbf{y}^{(t)} \rangle - 2\epsilon G \Omega_{\mathcal{X}} - \epsilon\eta, \quad (12)$$

for any  $\widehat{\mathbf{x}} \in \mathcal{X}$ . By symmetry, we analogously get that

$$\langle \mathbf{y}^{(t)}, \mathbf{B}^{\top} \mathbf{x}^{(t)} \rangle \geq \langle \widehat{\mathbf{y}}^{(t)}, \mathbf{B}^{\top} \mathbf{x}^{(t)} \rangle - \epsilon\eta \geq \langle \widehat{\mathbf{y}}, \mathbf{B}^{\top} \mathbf{x}^{(t)} \rangle - 2\epsilon G \Omega_{\mathcal{Y}} - \epsilon\eta, \quad (13)$$

for any  $\widehat{\mathbf{y}} \in \mathcal{Y}$ . Thus, recalling Definition 2.5, the claim follows from (12) and (13).  $\square$

**Lemma 3.3.** *Suppose that both players in a bimatrix game  $(\mathbf{A}, \mathbf{B})$  employ (OMD) with learning rate  $\eta > 0$ . Then, for any  $T \in \mathbb{N}$ ,*

$$\begin{aligned}\sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| &\geq \frac{1}{2\eta\|\mathcal{X}\|\|\mathbf{A}\|_{\text{op}}} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 + \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\|^2 \right) - \frac{2}{\|\mathbf{A}\|_{\text{op}}}; \\ \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\| &\geq \frac{1}{2\eta\|\mathcal{Y}\|\|\mathbf{B}\|_{\text{op}}} \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 + \|\widehat{\mathbf{y}}^{(t)} - \mathbf{y}^{(t)}\|^2 \right) - \frac{2}{\|\mathbf{B}\|_{\text{op}}}.\end{aligned}$$

*Proof.* By 1-strong convexity of  $\mathcal{R}_{\mathcal{X}}$  with respect to  $\|\cdot\|$ ,

$$\langle \mathbf{x}^{(t)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle - \frac{1}{\eta} D_{\mathcal{R}_{\mathcal{X}}}(\mathbf{x}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) - \langle \widehat{\mathbf{x}}^{(t-1)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle \geq \frac{1}{2\eta} \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2, \quad (14)$$

where we used the definition of the update rule of the primary sequence of (OMD). Similarly,

$$\langle \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle - \frac{1}{\eta} D_{\mathcal{R}_{\mathcal{X}}}(\widehat{\mathbf{x}}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) - \langle \mathbf{x}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle + \frac{1}{\eta} D_{\mathcal{R}_{\mathcal{X}}}(\mathbf{x}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) \geq \frac{1}{2\eta} \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\|^2. \quad (15)$$

Hence, summing (14) and (15) yields that

$$\langle \mathbf{x}^{(t)}, \mathbf{A}(\mathbf{y}^{(t-1)} - \mathbf{y}^{(t)}) \rangle \geq \frac{1}{2\eta} \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 + \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\|^2 \right) - \langle \widehat{\mathbf{x}}^{(t)}, \mathbf{A}\mathbf{y}^{(t)} \rangle + \langle \widehat{\mathbf{x}}^{(t-1)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle,$$

where we used that  $D_{\mathcal{R}_X}(\widehat{\mathbf{x}}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) \geq 0$ . Thus, a telescopic summation over all  $t \in [T]$  implies that

$$\sum_{t=1}^T \langle \mathbf{x}^{(t)}, \mathbf{A}(\mathbf{y}^{(t-1)} - \mathbf{y}^{(t)}) \rangle \geq \frac{1}{2\eta} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 + \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\|^2 \right) - 2\|\mathcal{X}\|, \quad (16)$$

since  $-\langle \widehat{\mathbf{x}}^{(T)}, \mathbf{A}\mathbf{y}^{(T)} \rangle \geq -\|\widehat{\mathbf{x}}^{(T)}\| \|\mathbf{A}\mathbf{y}^{(T)}\|_* \geq -\|\mathcal{X}\|$  and  $\langle \widehat{\mathbf{x}}^{(0)}, \mathbf{A}\mathbf{y}^{(0)} \rangle \geq -\|\widehat{\mathbf{x}}^{(0)}\| \|\mathbf{A}\mathbf{y}^{(0)}\|_* \geq -\|\mathcal{X}\|$ , where we used the normalization assumption. Furthermore,

$$\langle \mathbf{x}^{(t)}, \mathbf{A}(\mathbf{y}^{(t-1)} - \mathbf{y}^{(t)}) \rangle \leq \|\mathbf{x}^{(t)}\| \|\mathbf{A}(\mathbf{y}^{(t-1)} - \mathbf{y}^{(t)})\|_* \leq \|\mathcal{X}\| \|\mathbf{A}\|_{\text{op}} \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|.$$

Thus, combining this inequality with (16) implies that

$$\sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| \geq \frac{1}{2\eta \|\mathcal{X}\| \|\mathbf{A}\|_{\text{op}}} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 + \|\widehat{\mathbf{x}}^{(t)} - \mathbf{x}^{(t)}\|^2 \right) - \frac{2}{\|\mathbf{A}\|_{\text{op}}}.$$

This completes the first part of the claim. The second part follows analogously by symmetry.  $\square$

**Lemma 3.5** (Stability of OMD). *Suppose that both players employ (OMD) with learning rate  $\eta > 0$ . Then, for any  $t \in \mathbb{N}$ ,*

$$\begin{aligned} \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\| &\leq 3\eta; \\ \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| &\leq 3\eta. \end{aligned}$$

*Proof.* Fix any  $t \in \mathbb{N}$ . By definition of the primary sequence of (OMD),

$$\langle \mathbf{x}^{(t)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle - \frac{1}{\eta} D_{\mathcal{R}_X}(\mathbf{x}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) - \langle \widehat{\mathbf{x}}^{(t-1)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle \geq \frac{1}{2\eta} \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2,$$

where we used the 1-strong convexity of the regularizer  $\mathcal{R}_X$  with respect to  $\|\cdot\|$ . In turn, this implies that

$$\langle \mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}, \mathbf{A}\mathbf{y}^{(t-1)} \rangle \geq \frac{1}{\eta} \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2,$$

since  $D_{\mathcal{R}_X}(\mathbf{x}^{(t)} \parallel \widehat{\mathbf{x}}^{(t-1)}) \geq \frac{1}{2} \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2$  (by 1-strong convexity of  $\mathcal{R}_X$ ). Thus, an application of Cauchy-Schwarz inequality yields that

$$\|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \leq \eta \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \|\mathbf{A}\mathbf{y}^{(t-1)}\|_* \implies \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq \eta, \quad (17)$$

since  $\|\mathbf{A}\mathbf{y}^{(t-1)}\|_* \leq 1$  by the normalization assumption. Similar reasoning applied for the secondary sequence of (OMD) implies that for any  $t \in \mathbb{N}$ ,

$$\|\widehat{\mathbf{x}}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq \eta. \quad (18)$$

Now if  $t = 1$ , it follows from (17) that  $\|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\| = \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq \eta$  since  $\mathbf{x}^{(0)} = \widehat{\mathbf{x}}^{(0)}$ . Otherwise, for  $t \geq 2$ , applying the triangle inequality yields that  $\|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\| \leq \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| + \|\mathbf{x}^{(t-1)} - \widehat{\mathbf{x}}^{(t-2)}\| + \|\widehat{\mathbf{x}}^{(t-1)} - \widehat{\mathbf{x}}^{(t-2)}\| \leq 3\eta$  by (17) and (18). This completes the first part of the proof. Analogously, we conclude that  $\|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| \leq 3\eta$  for any  $t \in \mathbb{N}$ .  $\square$

**Theorem A.1** (Linear Decay of Regret; Full Version of Theorem 3.4). *Suppose that both players in a bimatrix game  $(\mathbf{A}, \mathbf{B})$  employ (OMD) with  $G$ -smooth regularizer, learning rate  $\eta > 0$  such that*

$$\eta \leq \min \left\{ \frac{1}{4 \max\{\|\mathbf{A}\|_{\text{op}}, \|\mathbf{B}\|_{\text{op}}\}}, \frac{\epsilon^2}{96 \|\mathbf{A}\|_{\text{op}} \|\mathbf{B}\|_{\text{op}} \max\{\|\mathcal{X}\|, \|\mathcal{Y}\|\}} \right\}$$

and

$$T \geq \max \left\{ \frac{16 \max\{\|\mathcal{X}\|, \|\mathcal{Y}\|\}}{\epsilon^2 \eta}, \frac{32 \max\{\Omega_{\mathcal{R}_X}, \Omega_{\mathcal{R}_Y}\}}{\epsilon^2 \eta^2}, \frac{2048 \max\{\Omega_{\mathcal{R}_Y} \|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2, \Omega_{\mathcal{R}_X} \|\mathcal{Y}\|^2 \|\mathbf{B}\|_{\text{op}}^2\}}{\epsilon^4 \eta^2} \right\},$$

for some fixed  $\epsilon > 0$ . Then, if the dynamics do not reach a  $(2\epsilon G \max\{\Omega_X, \Omega_Y\} + \epsilon\eta)$ -approximate NE, then

$$\max\{\text{Reg}_{\mathcal{X}}^T, \text{Reg}_{\mathcal{Y}}^T\} \leq -\min \left\{ \frac{\epsilon^2 \eta}{32}, \frac{\epsilon^4 \eta}{2048 \max\{\|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2, \|\mathcal{Y}\|^2 \|\mathbf{B}\|_{\text{op}}^2\}} \right\} T.$$

*Proof.* Suppose that there exists  $t \in \llbracket T \rrbracket$  such that

$$\left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) + \left( \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 \right) \leq \epsilon^2 \eta^2.$$

This would imply that  $\|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|, \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\| \leq \epsilon \eta$  and  $\|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t)}\|, \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\| \leq \epsilon \eta$ . In turn, by Proposition 3.2 it follows that the pair of strategies  $(\mathbf{x}^{(t)}, \mathbf{y}^{(t)})$  is a  $(2\epsilon G \max\{\Omega_{\mathcal{X}}, \Omega_{\mathcal{Y}}\} + \epsilon \eta)$ -approximate Nash equilibrium, contradicting our assumption. As a result, we conclude that for all  $t \in \llbracket T \rrbracket$ ,

$$\left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) + \left( \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 \right) \geq \epsilon^2 \eta^2.$$

Summing over all  $t \in \llbracket T \rrbracket$  yields that

$$\sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) + \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 \right) \geq \epsilon^2 \eta^2 T. \quad (19)$$

We distinguish between two cases. First, we treat the case where

$$\sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) \geq \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 \right). \quad (20)$$

Then, by virtue of (19),

$$\sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) \geq \frac{\epsilon^2 \eta^2}{2} T. \quad (21)$$

Further, by the triangle inequality and Young's inequality,

$$\|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \leq 2\|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 + 2\|\widehat{\mathbf{y}}^{(t-1)} - \mathbf{y}^{(t-1)}\|^2,$$

and summing over all  $t \in \llbracket T \rrbracket$  yields that

$$\begin{aligned} \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 &\leq 2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 + 2 \sum_{t=1}^T \|\widehat{\mathbf{y}}^{(t-1)} - \mathbf{y}^{(t-1)}\|^2 \\ &\leq 2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2 + 2 \sum_{t=1}^T \|\widehat{\mathbf{y}}^{(t)} - \mathbf{y}^{(t)}\|^2, \end{aligned}$$

where the last inequality follows since  $\widehat{\mathbf{y}}^{(0)} = \mathbf{y}^{(0)}$ . Hence, combining the latter bound with (20) implies that

$$\sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) \geq \frac{1}{2} \sum_{t=1}^T \|\mathbf{y}^{(t)} - \widehat{\mathbf{y}}^{(t-1)}\|^2. \quad (22)$$

Now we are ready to bound the regret of player  $\mathcal{X}$ . By Corollary 3.1,

$$\text{Reg}_{\mathcal{X}}^T \leq \frac{\Omega_{\mathcal{R}_X}}{\eta} + \eta \|\mathbf{A}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 - \frac{1}{4\eta} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right). \quad (23)$$

But (22) implies that

$$\begin{aligned} \eta \|\mathbf{A}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 - \frac{1}{8\eta} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \widehat{\mathbf{x}}^{(t-1)}\|^2 \right) \\ \leq \left( \eta \|\mathbf{A}\|_{\text{op}}^2 - \frac{1}{16\eta} \right) \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \leq 0, \end{aligned}$$

since  $\eta \leq \frac{1}{4\|\mathbf{A}\|_{\text{op}}}$ . From this we conclude that

$$\text{Reg}_{\mathcal{X}}^T \leq \frac{\Omega_{\mathcal{R}_X}}{\eta} - \frac{1}{8\eta} \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t-1)}\|^2 \right) \leq \frac{\Omega_{\mathcal{R}_X}}{\eta} - \frac{\epsilon^2 \eta}{16} T \leq -\frac{\epsilon^2 \eta}{32} T, \quad (24)$$

for  $T \geq \frac{32\Omega_{\mathcal{R}_X}}{\epsilon^2 \eta^2}$ , where we used (21). Next, we focus on the regret of player  $\mathcal{Y}$ . By (21) and Lemma 3.3,

$$\sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| \geq \frac{\epsilon^2 \eta}{4\|\mathcal{X}\|\|\mathbf{A}\|_{\text{op}}} T - \frac{2}{\|\mathbf{A}\|_{\text{op}}} \geq \frac{\epsilon^2 \eta}{8\|\mathcal{X}\|\|\mathbf{A}\|_{\text{op}}} T, \quad (25)$$

since  $T \geq \frac{16\|\mathcal{X}\|}{\epsilon^2 \eta}$ . Further, by Cauchy-Schwarz inequality,

$$\sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \geq \frac{1}{T} \left( \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\| \right)^2 \geq \frac{\epsilon^4 \eta^2}{64\|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2} T. \quad (26)$$

Now from Corollary 3.1, the regret of player  $\mathcal{Y}$  can be bounded as

$$\begin{aligned} \text{Reg}_{\mathcal{Y}}^T &\leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} + \eta \|\mathbf{B}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 - \frac{1}{4\eta} \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|^2 \right) \\ &\leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} + \eta \|\mathbf{B}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 - \frac{1}{8\eta} \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2, \end{aligned}$$

where we used that

$$\sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \leq 2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|^2 + 2 \sum_{t=1}^T \|\hat{\mathbf{y}}^{(t)} - \mathbf{y}^{(t)}\|^2.$$

Furthermore, by Lemma 3.5 and (26),

$$\eta \|\mathbf{B}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 \leq 9\eta^3 \|\mathbf{B}\|_{\text{op}}^2 T \leq \frac{\epsilon^4 \eta}{1024\|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2} T \leq \frac{1}{16\eta} \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2,$$

for  $\eta \leq \epsilon^2 (96\|\mathcal{X}\|\|\mathbf{A}\|_{\text{op}}\|\mathbf{B}\|_{\text{op}})^{-1}$ . As a result, (26) implies that

$$\text{Reg}_{\mathcal{Y}}^T \leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} - \frac{1}{16\eta} \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} - \frac{\epsilon^4 \eta}{1024\|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2} T \leq -\frac{\epsilon^4 \eta}{2048\|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2} T,$$

for  $T \geq \frac{2048\Omega_{\mathcal{R}_Y} \|\mathcal{X}\|^2 \|\mathbf{A}\|_{\text{op}}^2}{\epsilon^4 \eta^2}$ . Similarly, let us treat the case where

$$\sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|^2 \right) \geq \sum_{t=1}^T \left( \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t)}\|^2 + \|\mathbf{x}^{(t)} - \hat{\mathbf{x}}^{(t-1)}\|^2 \right).$$

Then, for  $\eta \leq \frac{1}{4\|\mathbf{B}\|_{\text{op}}}$  and  $T \geq \frac{32\Omega_{\mathcal{R}_Y}}{\epsilon^2 \eta^2}$ ,

$$\text{Reg}_{\mathcal{Y}}^T \leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} - \frac{1}{8\eta} \sum_{t=1}^T \left( \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t)}\|^2 + \|\mathbf{y}^{(t)} - \hat{\mathbf{y}}^{(t-1)}\|^2 \right) \leq \frac{\Omega_{\mathcal{R}_Y}}{\eta} - \frac{\epsilon^2 \eta}{16} T \leq -\frac{\epsilon^2 \eta}{32} T.$$

Moreover, for  $T \geq \frac{16\|\mathcal{Y}\|}{\epsilon^2 \eta}$ ,

$$\sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 \geq \frac{\epsilon^4 \eta^2}{64\|\mathcal{Y}\|^2 \|\mathbf{B}\|_{\text{op}}^2} T.$$

Thus, for  $\eta \leq \epsilon^2 (96\|\mathcal{Y}\|\|\mathbf{A}\|_{\text{op}}\|\mathbf{B}\|_{\text{op}})^{-1}$ ,

$$\eta \|\mathbf{A}\|_{\text{op}}^2 \sum_{t=1}^T \|\mathbf{y}^{(t)} - \mathbf{y}^{(t-1)}\|^2 \leq 9\eta^3 \|\mathbf{A}\|_{\text{op}}^2 T \leq \frac{\epsilon^4 \eta}{1024\|\mathcal{Y}\|^2 \|\mathbf{B}\|_{\text{op}}^2} T \leq \frac{1}{16\eta} \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2.$$

Finally, for  $T \geq \frac{2048\Omega_{\mathcal{R}_X}\|\mathcal{Y}\|^2\|\mathbf{B}\|_{\text{op}}^2}{\epsilon^4\eta^2}$ ,

$$\text{Reg}_{\mathcal{X}}^T \leq \frac{\Omega_{\mathcal{R}_X}}{\eta} - \frac{1}{16\eta} \sum_{t=1}^T \|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\|^2 \leq \frac{\Omega_{\mathcal{R}_X}}{\eta} - \frac{\epsilon^4\eta}{1024\|\mathcal{Y}\|^2\|\mathbf{B}\|_{\text{op}}^2} T \leq -\frac{\epsilon^4\eta}{2048\|\mathcal{Y}\|^2\|\mathbf{B}\|_{\text{op}}^2} T.$$

□

Next, we state the implication of Theorem A.1 in normal-form games under (OGD). In that setting, it holds that  $\|\mathcal{X}\|, \|\mathcal{Y}\| = 1$ ;  $\Omega_{\mathcal{X}}, \Omega_{\mathcal{Y}} \leq \sqrt{2}$ ;  $\Omega_{\mathcal{R}_X}, \Omega_{\mathcal{R}_Y} \leq 1$ ; and  $G = 1$ . Thus, we obtain the following simplified statement.

**Corollary A.2** (OGD in Normal-Form Games). *Suppose that both players in a bimatrix game  $(\mathbf{A}, \mathbf{B})$  employ (OGD) with learning rate  $\eta > 0$  such that*

$$\eta \leq \min \left\{ \frac{1}{4 \max\{\|\mathbf{A}\|_{\text{op}}, \|\mathbf{B}\|_{\text{op}}\}}, \frac{\epsilon^2}{96\|\mathbf{A}\|_{\text{op}}\|\mathbf{B}\|_{\text{op}}} \right\}$$

and

$$T \geq \max \left\{ \frac{16}{\epsilon^2\eta}, \frac{32}{\epsilon^2\eta^2}, \frac{2048 \max\{\|\mathbf{A}\|_{\text{op}}^2, \|\mathbf{B}\|_{\text{op}}^2\}}{\epsilon^4\eta^2} \right\},$$

for some fixed  $\epsilon > 0$ . Then,

- Either there exists  $t \in \llbracket T \rrbracket$  such that the pair of strategies  $(\mathbf{x}^{(t)}, \mathbf{y}^{(t)}) \in \mathcal{X} \times \mathcal{Y}$  constitutes an  $\epsilon(3 + \eta)$ -approximate Nash equilibrium;
- Or, otherwise, the average correlated distribution of play after  $T$  repetitions of the game is a

$$\min \left\{ \frac{\epsilon^2\eta}{32}, \frac{\epsilon^4\eta}{2048 \max\{\|\mathbf{A}\|_{\text{op}}^2, \|\mathbf{B}\|_{\text{op}}^2\}} \right\} - \text{strong coarse correlated equilibrium.}$$

Finally, we state an extension of Theorem 1.1 that establishes a dichotomy based on whether *most* of the iterates are approximate Nash equilibria—not just a *single* iterate. The proof is almost identical to the argument of Theorem A.1, and is therefore omitted.

**Corollary A.3.** *Suppose that both players in a bimatrix game employ (OGD) with learning rate  $\eta = O(\epsilon^2\delta)$  and  $T = \Omega\left(\frac{1}{\eta^2\epsilon^4\delta^2}\right)$  repetitions, for a sufficiently small  $\epsilon > 0$  and  $\delta \in (0, 1)$ . Then,*

- Either a  $1 - \delta$  fraction of the iterates is an  $\epsilon$ -approximate Nash equilibrium;
- Or, otherwise, the average correlated distribution of play is an  $\Omega(\epsilon^4\eta\delta^2)$ -strong CCE.

## B Description of the Game Instances

In this section we provide a detailed description of the game instances we used in our experiments in Section 4.2.

**Liar’s Dice** The first game we experimented on is *Liar’s dice*, a popular benchmark introduced by Lisý et al. [2015]. In our instantiation, each of the two players initially privately roles a *single* unbiased 4-face die. Then, the first player announces any face value up to 4, as well as the minimum number of dice the player believes have that value (among the dice of both players). Subsequently, each player in its own turn can either make a higher bid, or challenge the claim made by the previous player by declaring that player a “liar”. In particular, a bid is higher than the previous one if either the face value is higher, or if the claimed number of dices is greater. In case the current player challenges the previous bid, all dice have to be revealed. If the claim was valid, the last bidder wins and receives a reward of +1, while the challenger incurs a negative payoff of −1. Otherwise, the utilities obtained are reversed.

**Sheriff** Our second benchmark is a bargaining game inspired by the board game *Sheriff of Nottingham*, introduced by Farina et al. [2019b]. This game consists of two players: the *smuggler* and the *sheriff*. In our instantiation, the smuggler initially selects a number  $n \in \{0, 1, 2, 3, 4, 5\}$  which corresponds to the number of *illegal items* to be loaded in the cargo. Each illegal item has a fixed value of 1. Next, 2 rounds of bargaining between the two players follow. At each round, the smuggler decides on a *bribe* ranging from 0 to  $b := 3$  (inclusive), and the sheriff must decide whether or not the cargo will be inspected given the bribe amount. The sheriff’s decision is binding only in the *last round* of bargaining: if the sheriff accepts the bribe, the game stops with the smuggler obtaining a utility of  $n$  minus the bribe amount  $b$  proposed in the last bargaining round, while the sheriff receives a utility equal to  $b$ . In contrast, if the sheriff does not accept the bribe in last bargaining round and decides to inspect the cargo, there are two possible alternatives:

- If the cargo has no illegal items (*i.e.*  $n = 0$ ), the smuggler receives the fixed amount of 3, while sheriff incurs a negative payoff of  $-3$ ;
- Otherwise, the utility of the smuggler is set to  $-2n$ , while the utility of the Sheriff is  $2n$ .

**Battleship** Our next benchmark is *Battleship*, a parametric version of the popular board game introduced in [Farina et al., 2019b]. At the beginning, each player secretly places its ships on separate locations on a grid of size  $2 \times 2$ . Every ship has size 1 and a value of 4, and the placement is such that there is no overlap with any other ship. After the placement, players take turns at “firing” at their opponent’s ships. The game proceeds until either one player has sunk all of the opponent’s ships, or each player has completed  $r = 2$  rounds of firing. At the end of the game, each player’s payoff is the sum of the values of the opponent’s ships that were sunk, minus the sum of the values of the ships that the player has lost *multiplied by two*. The latter modification makes the game general-sum, and incentivizes players to be more risk-averse.

**Goofspiel** Our final benchmark is *Goofspiel*, introduced by Ross [1971]. In this game every player has a hand of cards numbered from 1 to  $h$ , where in our instantiation  $h := 3$ . An additional stack of  $h$  cards is shuffled and singled out as winning the current prize. In every turn a prize card is revealed, and players privately choose one of their cards to bid. The player with the highest card wins the current prize, while in case of a tie the prize card is discarded. Due to this tie-breaking mechanism, even two-player instances are general-sum. After the completion of  $h$  turns, players obtain the sum of the values of the prize cards they have won. Further, the instances we consider are of *limited information*—the actions of the other player are observed only at the end of the game. This makes the game strategically more involved as each player has less information about the opponent’s actions.