which shows that X's payoff in the stationary state is u^*, regardless of Y's strategy $\boldsymbol{y}$.

Below, we show that if X uses another strategy $\boldsymbol{x} \neq x^*\mathbf{1}$, there always is Y's strategy such that $v^{\mathrm{st}} > v^* \Leftrightarrow u^{\mathrm{st}} < u^*$. As X's non-equilibrium strategy, we assume the case $x_1 \neq x^*$ representatively. Then, Y's strategy $\boldsymbol{y} = y^*\mathbf{1} + \mathrm{d}y_1\boldsymbol{e}_1$ with sufficiently small $\mathrm{d}y_1$ satisfies

$$\boldsymbol{p}^{\mathrm{st}} = \begin{pmatrix} x_1(y^* + \mathrm{d}y_1) & x_2y^* & x_3y^* & x_4y^* \\ x_1(\tilde{y}^* - \mathrm{d}y_1) & x_2\tilde{y}^* & x_3\tilde{y}^* & x_4\tilde{y}^* \\ \tilde{x}_1(y^* + \mathrm{d}y_1) & \tilde{x}_2y^* & \tilde{x}_3y^* & \tilde{x}_4y^* \\ \tilde{x}_1(\tilde{y}^* - \mathrm{d}y_1) & \tilde{x}_2\tilde{y}^* & \tilde{x}_3\tilde{y}^* & \tilde{x}_4\tilde{y}^* \end{pmatrix} \boldsymbol{p}^{\mathrm{st}}. \tag{A27}$$

In this equation, we approximate $\boldsymbol{p}^{\mathrm{st}} \simeq \boldsymbol{p}^{\mathrm{st}(0)} + \boldsymbol{p}^{\mathrm{st}(1)}$, where $\boldsymbol{p}^{\mathrm{st}(k)}$ describes the $O((\mathrm{d}y_1)^k)$ term in $\boldsymbol{p}^{\mathrm{st}}$. We can derive these 0-th and 1-st order terms by comparing the left-hand and right-side of this equation. Here, the 0-th order term satisfies $\boldsymbol{p}^{\mathrm{st}(0)} \cdot \boldsymbol{u} = u^*$, which means that the term does not contribute to the deviation from the Nash equilibrium payoff. On the other hand, the 1-st order term gives

$$\boldsymbol{p}^{\mathrm{st}(1)} = p_1^{\mathrm{st}(0)}\mathrm{d}y_1(+x_1, -x_1, +\tilde{x}_1, -\tilde{x}_1)^{\mathrm{T}} \tag{A28}$$

$$\Rightarrow v^{\mathrm{st}(1)} = p_1^{\mathrm{st}(0)}\mathrm{d}y_1 \underbrace{(v_1 - v_2 - v_3 + v_4)}_{=\boldsymbol{v}\cdot\mathbf{1}_z \neq 0}(x_1 - x^*). \tag{A29}$$

Here, we use $\mathbf{1}_z := (+1, -1, -1, +1)$. Thus, in the leading order, $v^{\mathrm{st}(1)} > v^* \Leftrightarrow u^{\mathrm{st}(1)} < u^*$ holds by taking $\mathrm{d}y_1 > 0$ if $\boldsymbol{v} \cdot \mathbf{1}_z(x_1 - x^*) > 0$, while by taking $\mathrm{d}y_1 < 0$ if $\boldsymbol{v} \cdot \mathbf{1}_z(x_1 - x^*) < 0$. In other words, X's minimax strategy is $\boldsymbol{x} = x^*\mathbf{1}$. Similarly, we can prove that Y's minimax strategy is $\boldsymbol{y} = y^*\mathbf{1}$. Thus, the Nash equilibrium is given by $(\boldsymbol{x}, \boldsymbol{y}) = (x^*\mathbf{1}, y^*\mathbf{1})$. □

B Analysis of Learning Dynamics

B.1 Simpler MMGA for Two-action Games

This section is concerned with the contents in **Section 4.2** in the main manuscript.

Especially in two-action games, we can use the formulation of Assumption 1 in the main manuscript. By replacing the strategies $(\mathbf{x}, \mathbf{y})$ by $(\boldsymbol{x}, \boldsymbol{y})$, we can formulation another simpler algorithm of MMGA as

Algorithm A1 Discretized MMGA for two-action

Input: η, γ

1: **for** $t = 0, 1, 2, \cdots$ **do**
2: **for** $i = 1, 2, \ldots, |\mathcal{S}|$ **do**
3: $\boldsymbol{x}' \leftarrow \boldsymbol{x}$
4: $x'_i \leftarrow x'_i + \gamma$
5: $\Delta_i \leftarrow (1 - x_i)\dfrac{u^{\mathrm{st}}(\boldsymbol{x}', \boldsymbol{y}) - u^{\mathrm{st}}(\boldsymbol{x}, \boldsymbol{y})}{\gamma}$
6: **end for**
7: **for** $i = 1, 2, \ldots, |\mathcal{S}|$ **do**
8: $x_i \leftarrow x_i(1 + \eta\Delta_i)$
9: **end for**
10: $\mathbf{x} \leftarrow \mathrm{Norm}(\mathbf{x})$
11: **end for**

There is a major difference between the original and simpler MMGAs in lines 4 and 5. The equivalence