Instance-dependent Sample Complexity for Bilinear Saddle-Point Optimization with Noisy Feedback: An LP-Based Approach

Jiashuo Jiang
Department of Industrial Engineering and Decision Analytics
Hong Kong University of Science and Technology
jsjiang@ust.hk

Mengxiao Zhang
Department of Business Analytics
Tippie College of Business, University of Iowa
mengxiao-zhang@uiowa.edu

Abstract

In this work, we study the sample complexity of obtaining a Nash equilibrium (NE) estimate in two-player zero-sum matrix games with noisy feedback. Specifically, we propose a novel algorithm that repeatedly solves linear programs (LPs) to obtain an NE estimate with bias at most ε with a sample complexity of $\mathcal{O}\left(\frac{m_1 m_2}{\varepsilon \min\{\delta^2, \sigma_0^2, \sigma^3\}} \log \frac{m_1 m_2}{\varepsilon}\right)$ for general $m_1 \times m_2$ game matrices, where σ, σ_0, δ are some problem-dependent constants. To our knowledge, this is the first instance-dependent sample complexity bound for finding an NE estimate with ε bias in general-dimension matrix games with noisy feedback and potentially non-unique equilibria. Our algorithm builds on recent advances in online resource allocation and operates in two stages: (1) identifying the support set of an NE, and (2) computing the unique NE restricted to this support. Both stages rely on a careful analysis of LP solutions derived from noisy samples.

1 Introduction

Minimax optimization [39] is a fundamental problem in machine learning and game theory, with applications including adversarial training [19] and robust optimization [36]. A canonical example of a minimax problem is the two-player zero-sum matrix game, also known as the bilinear saddle-point problem, defined as $\min_{\mathbf{x}} \max_{\mathbf{y}} \mathbf{x}^\top A \mathbf{y}$, where A is the payoff matrix, and $(\mathbf{x}, \mathbf{y})$ are the strategies of the two players lying within the simplex. The goal is to compute a *Nash Equilibrium (NE)*, a strategy pair in which neither player can improve their expected payoff by deviating from their strategy.

A significant line of research focuses on computing the NE when the game matrix A is fixed and fully observed. However, much less is known about computing NE when the learner only receives noisy and bandit feedback, when each player picks an action and only receives a noisy observation of A on the chosen action pair. While it is possible to find an ϵ-approximate NE using $\mathcal{O}(1/\epsilon^2)$ samples by sampling each action pair $\mathcal{O}(1/\epsilon^2)$ times and computing NE based on the empirical utility matrix, it remains unclear whether a better instance-dependent sample complexity can be achieved. Therefore, this paper aims to investigate the following question:

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

How to achieve better instance-dependent sample complexity for computing a Nash Equilibrium in two-player zero-sum matrix games with noisy and bandit feedback?

Problem Formulation. We consider the bilinear saddle point problem defined as follows

$$\min_{\mathbf{x}\in\mathcal{X}=\Delta_{m_1}} \max_{\mathbf{y}\in\mathcal{Y}=\Delta_{m_2}} \mathbf{x}^\top A\mathbf{y}, \tag{1}$$

where $\Delta_n = \{\mathbf{x}\in\mathbb{R}^n : \sum_{i=1}^n \mathbf{x}_i = 1, \mathbf{x}_i \geq 0, \forall i \in [n]\}$ denotes the $(n-1)$-dimensional simplex and $A \in [-1,1]^{m_1\times m_2}$ is the payoff matrix. According to the celebrated minimax theorem [39], we know that $\min_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top A\mathbf{y} = \max_{\mathbf{y}\in\mathcal{Y}} \min_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top A\mathbf{y}$. Define $\mathcal{X}^*\times\mathcal{Y}^*$ as the set of Nash Equilibria where $\mathcal{X}^* = \operatorname{argmin}_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top A\mathbf{y}$ and $\mathcal{Y}^* = \operatorname{argmax}_{\mathbf{y}\in\mathcal{Y}} \min_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top A\mathbf{y}$. In the following, we introduce two definitions that measure of the closeness between a strategy (x,y) and the Nash Equilibria set $\mathcal{X}^*\times\mathcal{Y}^*$.

Definition 1 *We call a pair of strategy* $(\mathbf{x},\mathbf{y}) \in \Delta_{m_1}\times\Delta_{m_2}$ *an* ϵ-close NE *if* $\operatorname{argmin}_{\mathbf{x}^*\in\mathcal{X}^*} \|\mathbf{x}-\mathbf{x}^*\|_2 \leq \epsilon$ *and* $\operatorname{argmin}_{\mathbf{y}^*\in\mathcal{Y}} \|\mathbf{y}-\mathbf{y}^*\|_2 \leq \epsilon$*.*

Definition 2 *Define the sub-optimality gap of* $x \in \mathcal{X}$ *and* $y \in \mathcal{Y}$ *as* $\max_{\mathbf{y}'\in\Delta_{m_2}} \mathbf{x}^\top A\mathbf{y}' - \max_{\mathbf{y}'\in\Delta_{m_2}} x^{*\top} A\mathbf{y}'$ *and* $\min_{\mathbf{x}'\in\Delta_{m_1}} \mathbf{x}'^\top A\mathbf{y}^* - \min_{\mathbf{x}'\in\Delta_{m_1}} x'^\top A\mathbf{y}$ *(independent fo the choice of* $x^*\in\mathcal{X}^*$ *and* $y^*\in\mathcal{Y}$*). We call* $x\in\mathcal{X}$ *(*$y\in\mathcal{Y}$*) an* ϵ*-sub-optimality-gap NE if* x*'s (*y*'s) sub-optimality gap is no more than* ϵ*.*

Since we assume that $A\in[-1,1]^{m_1\times m_2}$, we know that an ϵ-close NE implies an ϵ-sub-optimality-gap NE, while the reverse is not true. In this paper, we consider the setting where only noisy bandit feedback is available. Specifically, the learner can only query an entry (i,j) of A and observe a sample $A_{i,j}+\eta\in[-1,1]$ where η is a zero-mean noise.

Main Results and Contributions. Inspired by the recent advance of online resource allocation [40, 29, 25, 30], we design an algorithm based on resolving linear programs (LP) achieving an ϵ-approximate NE with sample complexity $\mathcal{O}(\frac{m}{\epsilon\min\{\delta^2,\sigma_0^2,\sigma^3\}}\log(\frac{m}{\epsilon}))$, where $m = m_1m_2$ and δ, σ_0 and σ are instance-dependent values related to LP.[1] Specifically, our algorithm includes two components.

The first phase is to identify the support of an NE. Specifically, we first sample each entry of matrix N_1/m times and construct an empirical LP formulation of the game. Our algorithm then iteratively computes the estimated support of the NE by restricting the strategy to progressively smaller supports. Concretely, at each step, we select an entry from the current support and compare the objective value of the empirical LP in two cases: one where the strategy remains unrestricted and one where the selected entry is removed. Otherwise, we terminate the process and output the current support. With a high probability, this guarantees to find the support of an NE when N_1 exceeds certain instance-dependent constant.

After obtaining an optimal basis of an NE from the first phase, the second component is to find the NE on this support. Specifically, inspired by the recent advances in solving online resource allocation that achieves logarithmic regret via LP resolving (e.g. [40, 29, 25, 30]), our algorithm iteratively solves the LP based on the historical samples with an adaptive choice of the constraint slackness. We show that our algorithm achieves an instance-dependent sample complexity of $\mathcal{O}(\frac{m}{\delta^2\epsilon}\log\frac{m}{\delta\epsilon})$, then complementing with a robustness guarantee $\mathcal{O}(\frac{m}{\epsilon^2}\log\frac{m}{\epsilon})$, which does not depend on the instance-dependent value δ.[2]

2 Our Approach

Reformulation as Linear Programming In order to solve Equation (1), we can focus on two linear programs, which can be regarded as the primal-dual form to each other. We have the following

[1]All problem-dependent constants are formally defined in later sections.

[2]We omit some other problem-dependent constants for conciseness. These dependencies are explicitly written in the theorem statement.

"primal" linear program.

$$V^{\text{Prime}} = \min_{\mathbf{x}\succeq 0, \mu\in\mathbb{R}} \mu \quad \text{s.t. } \mu \cdot \mathbf{e}^{m_2} \succeq A^\top \mathbf{x},\ (\mathbf{e}^{m_1})^\top \mathbf{x} = 1. \tag{2}$$

According to the celebrated minimax theorem, Equation (1) is equivalent to $\max_{\mathbf{y}\in\mathcal{Y}} \min_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top A\mathbf{y}$. We also have the "dual" linear programming that computes $\mathbf{y}$.

$$V^{\text{Dual}} = \max_{\nu\geq 0, \mathbf{y}\succeq 0} \nu \quad \text{s.t. } \nu \cdot \mathbf{e}^{m_1} \preceq A\mathbf{y},\ (\mathbf{e}^{m_2})^\top \mathbf{y} = 1. \tag{3}$$

Classic LP analysis (e.g. Theorem 16.5 in [18]) shows that it suffices to find an optimal solution $\mathbf{x}^*$ to LP Equation (2) and an optimal solution $\mathbf{y}^*$ to LP Equation (3), such that $\mathbf{x}^*$ and $\mathbf{y}^*$ are the corresponding primal-dual solutions.

Saddle Point Support Identification. In order to identify the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$ from noisy observations, we first sample each entry $A_{i,j}$ a number of times and construct an empirical game matrix $\widehat{A}$ with each entry the empirical average of the N observations, and construct the primal LP of Equation (2) using $\widehat{A}$. Next, we successively shrink the size of x-player's support based on the objective value $\widehat{V}_{\mathcal{I}}^{\text{Prime}}$. Specifically, $\widehat{V}_{\mathcal{I}}^{\text{Prime}}$ is defined as the empirical primal LP with non-zero values only on $\mathcal{I}$ and we initialize the support set to be $\mathcal{I} = [m_1]$. At each time, we sequentially pick each $i \in \mathcal{I}$ and compare the objective value of $\widehat{V}_{\mathcal{I}}^{\text{Prime}}$ with $\widehat{V}_{\mathcal{I}\backslash\{i\}}^{\text{Prime}}$. If these two quantities are the same, this means that there exists an NE that is not supported on i and we set $\mathcal{I} \leftarrow \mathcal{I}\backslash\{i\}$; otherwise, we keep i in $\mathcal{I}$. We then obtain the support set $\mathcal{I}^*$ after iterating over all $i \in [m_1]$.

After identifying the set $\mathcal{I}^*$, we also need to find the set $\mathcal{J}^*$, which is equivalent to see what is the support for the dual LP. We apply a similar approach to find $\mathcal{J}^*$. Specifically, based on $\mathcal{I}^*$, we define the dual LP $\widehat{V}_{\mathcal{I}^*,\mathcal{J}}^{\text{Dual}}$ in which we only include the rows of $\widehat{A}$ whose indices are in $\mathcal{I}^*$ and set the dual variables outside $\mathcal{J}$ to be 0. We initialize $\mathcal{J} = [m_2]$ and successively eliminate an element $j \in \mathcal{J}$ if $\widehat{V}_{\mathcal{I}^*,\mathcal{J}\backslash\{j\}}$ equals to $\widehat{V}_{\mathcal{I}^*,\mathcal{J}}^{\text{Prime}}$. Finally, we obtain $\mathcal{J}^*$ after iterating over all $j \in [m_2]$. Note that it is possible that there are multiple optimal basis of $\mathcal{I}^*$ and $\mathcal{J}^*$ and we show later that our approach essentially identifies one particular optimal basis, since our goal is to find one NE of the game. In the following, we provide both instance-dependent and worst-case sample complexity guarantees.

LP Resolving Procedure. After obtaining the support $\mathcal{I}^*$ and $\mathcal{J}^*$, we conduct LP resolving to approximate the optimal solution. Note that the optimal solution $\mathbf{x}^*$ enjoys the following structure: $\mathbf{x}^*_{\mathcal{I}^{*c}} = 0$ and the other elements $\mathbf{x}^*_{\mathcal{I}^*}$ can be given as the solution to

$$\begin{bmatrix} A^\top_{\mathcal{I}^*,\mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}^*_{\mathcal{I}^*} \\ \mu^* \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{4}$$

Our approach is formalized in Algorithm 1. Specifically, in each iteration n, we observe a noisy matrix entry $\tilde{A}_{n,i_n,j_n}$ with $\mathbb{E}[\tilde{A}_{n,i_n,j_n}] = A_{i_n,j_n}$ and (i_n, j_n) uniformly drawn from $\mathcal{I}^* \times \mathcal{J}^*$. A naive approach would be to directly use these samples to form an empirical estimate of $A_{\mathcal{I}^*,\mathcal{J}^*}$ and compute the optimal basis. However, this would result in estimation error of $\Theta(1/\sqrt{N})$ by standard concentration bounds, leading to a suboptimal $\Theta(1/\sqrt{N})$ approximate Nash Equilibrium.

To achieve a better convergence rate, Algorithm 1 is inspired by the LP resolving algorithms for online resource allocation problems [40, 29, 25, 30] that can achieve $\mathcal{O}(\log N)$ regret after N iterations. Specifically, we interpret Equation (4) as a resource allocation problem: the right-hand-side represents resources, variables $\mathbf{x}$ represent actions, and the left-hand-side matrix represents resource consumption. The goal is to determine actions $\{\mathbf{x}_n\}_{n\in[N]}$ such that resources are appropriately consumed over N iterations. The crucial point of Algorithm 2's effectiveness lies in its adaptive resource allocation procedure (Line 8 of Algorithm 1). By dynamically updating $\mathbf{a}^n$ according to Equation (6), we create a self-correcting system that forms and solves the LP for $\tilde{\mathbf{x}}^n$ as specified in Equation (5). This resolving procedure ensures that $\mathbf{a}^n$ converges to $\mathbf{a}$ with a gap bounded by $O(1/(N-n+1))$ – a critical factor in achieving our instance-dependent guarantee. For example, when a binding constraint $j \in \mathcal{J}^*$ has a value below its target $\mathbf{a}_j$ under the current action $\sum_{n'=1}^{n} \mathbf{x}^{n'}/n$, the algorithm compensates by setting $\mathbf{a}_j^n/(N-n+1)$ greater than $\mathbf{a}_j$, thereby increasing the constraint value for $\tilde{\mathbf{x}}^{n+1}$. Through this continuous feedback loop, the algorithm dynamically adjusts for constraint violations, enabling the significantly fast convergence rate of $O(1/(N-n+1))$.

Algorithm 1 The Resolving Algorithm

Input: the number of samples N, the support $\mathcal{I}^*$ and $\mathcal{J}^*$ with $|\mathcal{I}^*| = |\mathcal{J}^*|$, $L = 4$.
Initialize $\mathcal{H}^1 = \emptyset$, $\mathbf{a}^1 = \mathbf{0} \in \mathbb{R}^{|\mathcal{J}^*|}$
for $n = 1, \ldots, N$ **do**
    Construct estimates $\widehat{A}_{\mathcal{I}^*, \mathcal{J}^*}$ using the dataset $\mathcal{H}^n$.
    Construct a solution $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ such that $\tilde{\mathbf{x}}^n_{\mathcal{I}^{*c}} = 0$ and $\tilde{\mathbf{x}}^n_{\mathcal{I}^*}$ is the solution to

$$\begin{bmatrix} \widehat{A}^\top_{\mathcal{I}^*, \mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_{\mathcal{I}^*} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{a}^n / (N - n + 1) \\ 1 \end{bmatrix}. \tag{5}$$

    Project $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ to the set $\{(\mathbf{x}, \mu) : \mathbf{x} \geq 0, \|(\mathbf{x}, \mu)\|_1 \leq L\}$ to obtain $(\mathbf{x}^n, \mu^n)$.
    Observe $\tilde{A}_{n, i_n, j_n} = A_{i_n, j_n} + \eta_n$, where i_n and j_n are uniformly sampled from $\mathcal{I}^*$ and $\mathcal{J}^*$.
    Update $\mathcal{H}^{n+1} = \mathcal{H}^n \cup \{(i_n, j_n, \tilde{A}_{n, i_n, j_n})\}$ and compute $\mathbf{a}^{n+1}$ as

$$\mathbf{a}^{n+1} = \mathbf{a}^n - |\mathcal{J}^*| \cdot |\mathcal{I}^*| \cdot \tilde{A}_{i_n, j_n} \cdot x^n_{i_n} \cdot \mathbf{h}_{j_n} + \mu^n \cdot \mathbf{e}^{|\mathcal{J}^*|}, \tag{6}$$

    where $\mathbf{h}_{j_n} \in \mathbb{R}^{|\mathcal{J}^*|}$ denotes a vector with j_n-th position being 1 and other positions being 0.
Compute $\bar{\mathbf{x}} = \frac{1}{N} \cdot \sum_{n=1}^N \mathbf{x}^n$ and output $\bar{\mathbf{x}}$.

3 Sample Complexity Analysis

We now show that our Algorithm 1 achieves an instance-dependent sample complexity for approximating the optimal solution to the primal LP problem V^{Prime}.

Theorem 1 *For any $\epsilon > 0$, with a sample complexity of $N = \mathcal{O}\left(\left(\frac{d^{5/2}}{\sigma^3} + \frac{m}{\delta^2}\right) \cdot \frac{\log(m/\varepsilon)}{\varepsilon}\right)$, where δ being certain problem-dependent constant and $d = |\mathcal{I}^*| = |\mathcal{J}^*|$, applying Algorithm 1 guarantees that* $\operatorname{argmin}_{\mathbf{x}^* \in \mathcal{X}^*} \|\mathbb{E}[\bar{\mathbf{x}}] - \mathbf{x}^*\|_2 \leq \varepsilon$. *Here, $\sigma > 0$ is the minimum singular value of the full-rank matrix A^* defined as*

$$A^* = \begin{bmatrix} A^\top_{\mathcal{I}^*, \mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top, & 0 \end{bmatrix}. \tag{7}$$

we show that Algorithm 1 also achieves an δ-independent sub-optimality gap with N samples, demonstrating the robustness of our algorithms. Specifically, we have the following theorem:

Theorem 2 *For a fixed sample size N, denote by $\bar{\mathbf{x}}$ the output of Algorithm 1. We know that $\mathbb{E}[\bar{\mathbf{x}}]$ is a feasible solution to V^{Prime} with an suboptimality gap upper bounded by*

$$\mathcal{O}\left(\frac{d^2}{\sigma} \cdot \left(1 + \frac{d}{\sigma^2 \cdot L^2}\right) \cdot \frac{\log(N)}{N} + \widehat{\kappa} \cdot \sqrt{\frac{dm \cdot \log(N)}{N}}\right),$$

where $\widehat{\kappa}$ is certain constant independent of N, $d = |\mathcal{I}^*| = |\mathcal{J}^*|$, and σ is the minimum singular value of A^* defined in Theorem 1.

Combining the above two thoerems, we have the following result.

Theorem 3 *For a fixed accuracy level $\varepsilon > 0$, denote by $\bar{\mathbf{x}}$ the output of Algorithm 1 with N samples. We know that $\mathbb{E}[\bar{\mathbf{x}}]$ forms a feasible solution to V^{Prime}, with an suboptimality gap upper bounded by ε, as long as the sample complexity bound holds*

$$N = \mathcal{O}\left(\frac{d^3}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + m \cdot \min\left\{\frac{d \cdot \widehat{\kappa}^2}{\varepsilon^2}, \frac{1}{\delta^2 \cdot \varepsilon}\right\} \cdot \log(m/\varepsilon)\right),$$

where δ being certain problem-dependent constant, and σ the minimum singular value of A^* defined in Theorem 1, and $\widehat{\kappa}$ is certain constant independent of ε.

References

[1] S. Agrawal, Z. Wang, and Y. Ye. A dynamic near-optimal algorithm for online linear programming. *Operations Research*, 62(4):876–890, 2014.

[2] I. Anagnostides, I. Panageas, G. Farina, and T. Sandholm. On last-iterate convergence beyond zero-sum games. In *International Conference on Machine Learning*, pages 536–581. PMLR, 2022.

[3] R. Ao, J. Jiang, and D. Simchi-Levi. Learning to price with resource constraints: From full information to machine-learned prices. *arXiv preprint arXiv:2501.14155*, 2025.

[4] A. Arlotto and X. Xie. Logarithmic regret in the dynamic and stochastic knapsack problem with equal rewards. *Stochastic Systems*, 10(2):170–191, 2020.

[5] S. Banerjee and D. Freund. Good prophets know when the end is near. *Management Science*, 2024.

[6] N. Buchbinder and J. Naor. Online primal-dual algorithms for covering and packing. *Mathematics of Operations Research*, 34(2):270–286, 2009.

[7] P. Bumpensanti and H. Wang. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science*, 66(7):2993–3009, 2020.

[8] Y. Cai, G. Farina, J. Grand-Clément, C. Kroer, C.-W. Lee, H. Luo, and W. Zheng. Fast last-iterate convergence of learning in games requires forgetful algorithms. *arXiv preprint arXiv:2406.10631*, 2024.

[9] Y. Cai, A. Oikonomou, and W. Zheng. Tight last-iterate convergence of the extragradient and the optimistic gradient descent-ascent algorithm for constrained monotone variational inequalities. *arXiv preprint arXiv:2204.09228*, 2022.

[10] J. Dallant, F. Haagensen, R. Jacob, L. Kozma, and S. Wild. Finding the saddlepoint faster than sorting. In *2024 Symposium on Simplicity in Algorithms (SOSA)*, pages 168–178. SIAM, 2024.

[11] J. Dallant, F. Haagensen, R. Jacob, L. Kozma, and S. Wild. An optimal randomized algorithm for finding the saddlepoint. *arXiv preprint arXiv:2401.06512*, 2024.

[12] C. Daskalakis, A. Deckelbaum, and A. Kim. Near-optimal no-regret algorithms for zero-sum games. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 235–254. SIAM, 2011.

[13] C. Daskalakis and I. Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. *arXiv preprint arXiv:1807.04252*, 2018.

[14] R. S. Falk. Lecture note 3, perturbation theory for linear systems of equations. 2016.

[15] T. S. Ferguson. Who solved the secretary problem? *Statistical science*, 4(3):282–289, 1989.

[16] Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

[17] G. Gallego and G. Van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations research*, 45(1):24–41, 1997.

[18] S. O. Gharan. Lecture note 16, Duality and the Minimax theorem. 2016.

[19] I. J. Goodfellow, J. Shlens, and C. Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.

[20] A. Gupta and M. Molinaro. How experts can solve lps online. In *European Symposium on Algorithms*, pages 517–529. Springer, 2014.

[21] Y.-G. Hsieh, K. Antonakopoulos, and P. Mertikopoulos. Adaptive learning in continuous games: Optimal regret bounds and convergence to nash equilibrium. In *Conference on Learning Theory*, pages 2388–2422. PMLR, 2021.

[22] S. Ito, H. Luo, T. Tsuchiya, and Y. Wu. Instance-dependent regret bounds for learning two-player zero-sum games with bandit feedback. *Conference on Learning Theory*, 2025.

[23] S. Jasin. Reoptimization and self-adjusting price control for network revenue management. *Operations Research*, 62(5):1168–1178, 2014.

[24] S. Jasin and S. Kumar. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research*, 37(2):313–345, 2012.

[25] J. Jiang, W. Ma, and J. Zhang. Degeneracy is ok: Logarithmic regret for network revenue management with indiscrete distributions. *Operations Research*, 2025.

[26] S. Kalyanakrishnan, A. Tewari, P. Auer, and P. Stone. Pac subset selection in stochastic multi-armed bandits. In *ICML*, volume 12, pages 655–662, 2012.

[27] T. Kesselheim, A. Tönnis, K. Radke, and B. Vöcking. Primal beats dual on online packing lps in the random-order model. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 303–312, 2014.

[28] X. Li, C. Sun, and Y. Ye. The symmetry between arms and knapsacks: A primal-dual approach for bandits with knapsacks. In *International Conference on Machine Learning*, pages 6483–6492. PMLR, 2021.

[29] X. Li and Y. Ye. Online linear programming: Dual convergence, new algorithms, and regret bounds. *Operations Research*, 70(5):2948–2966, 2022.

[30] W. Ma, Y. Cao, D. H. Tsang, and D. Xia. Optimal regularized online allocation by adaptive re-solving. *Operations Research*, 2025.

[31] A. Maiti, R. Boczar, K. Jamieson, and L. Ratliff. Near-optimal pure exploration in matrix games: A generalization of stochastic bandits & dueling bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 2602–2610. PMLR, 2024.

[32] A. Maiti, R. Boczar, K. Jamieson, and L. J. Ratliff. Query-efficient algorithms to find the unique nash equilibrium in a two-player zero-sum matrix game. *arXiv preprint arXiv:2310.16236*, 2023.

[33] A. Maiti, K. Jamieson, and L. J. Ratliff. Logarithmic regret for matrix games against an adversary with noisy bandit feedback. *arXiv preprint arXiv:2306.13233*, 2023.

[34] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22–es, 2007.

[35] M. Molinaro and R. Ravi. The geometry of online packing linear programs. *Mathematics of Operations Research*, 39(1):46–59, 2014.

[36] H. Rahimian and S. Mehrotra. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*, 2019.

[37] S. Rakhlin and K. Sridharan. Optimization, learning, and games with predictable sequences. *Advances in Neural Information Processing Systems*, 26, 2013.

[38] V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire. Fast convergence of regularized learning in games. *Advances in Neural Information Processing Systems*, 28, 2015.

[39] J. v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.

[40] A. Vera and S. Banerjee. The bayesian prophet: A low-regret framework for online decision making. *Management Science*, 67(3):1368–1391, 2021.

[41] C.-Y. Wei, C.-W. Lee, M. Zhang, and H. Luo. Linear last-iterate convergence in constrained saddle-point optimization. *arXiv preprint arXiv:2006.09517*, 2020.

[42] Y. Wei, J. Xu, and S. H. Yu. Constant regret primal-dual policy for multi-way dynamic matching. In *Abstract Proceedings of the 2023 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, pages 79–80, 2023.

[43] H. Weyl. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen*, 71(4):441–479, 1912.

[44] Y. Xie, W. Ma, and L. Xin. The benefits of delay to online decision-making. *Management Science*, 2025.

[45] Y. Zhou, J. Li, and J. Zhu. Identify the nash equilibrium in static games with random payoffs. In *International Conference on Machine Learning*, pages 4160–4169. PMLR, 2017.

[46] J. Zimmert and Y. Seldin. Tsallis-inf: An optimal algorithm for stochastic and adversarial bandits. *Journal of Machine Learning Research*, 22(28):1–49, 2021.

A Related Works

There is a line of works considering the sample complexity of finding NE for matrix games. When the observations are deterministic, [32] proposed an inefficient algorithm that identifies all Nash Equilibria by querying $\mathcal{O}(nk^5 \log^2 n)$ entries of the matrix, where k denotes the size of the support across all Nash Equilibria and $A \in \mathbb{R}^{n\times n}$. For Pure Strategy Nash Equilibria (PSNE), a special class of NE where each player's strategy is restricted to a single action, [10] developed a deterministic and efficient algorithm requiring only $\mathcal{O}(n)$ samples. This was subsequently improved by [11], who introduced a randomized algorithm that further reduced the runtime complexity while maintaining the same sample complexity.

When the observations are noisy, [45] proposed a lower-upper-confidence-bound (LUCB)-based algorithm [26] for identifying pure-strategy Nash equilibria (PSNE) with probability at least $1-\delta$, achieving a sample complexity of $\mathcal{O}(\log(1/\delta))$. [31] later improved this result by refining problem-dependent constants to achieve better sample efficiency. Similar to our approach, their algorithm consists of two subroutines: the first identifies the support for the x-player, with a size of at most 2, and the second computes the unique NE on the resulting 2×2 sub-matrix. [33] considered the case where one player competes against an adversary, deriving logarithmic Nash regret guarantees.

There is another related line of work which aims to find the NE of a matrix game via regret minimization. Specifically, [16] shows that if both players adopt an algorithm with individual regret bound R over T rounds, then the average-iterate for each player forms an $\frac{R}{T}$-approximate NE. Since $R = \Theta(\sqrt{T})$ in the worst case, this leads to an $\mathcal{O}(1/\sqrt{T})$-approximate NE. Later, a line of works [12, 37, 38] show that with full gradient feedback, if both players apply certain optimistic online learning algorithms, the average-iterate converges to NE with a faster rate. A recent work [22] shows that under noisy bandit feedback, if both players apply Tsallis-INF [46], then the expected average-iterate converges to an $\widetilde{\mathcal{O}}(1/\sqrt{T})$-Nash Equilibrium, with a faster $\mathcal{O}(\log T/T)$ instance-dependent rate if the game has a unique PSNE. Besides average-iterate convergence, there are also works studying the last-iterate convergence [13, 41, 21, 2, 9, 8, 22], which is more preferable in many applications.

Our algorithm design is inspired by the recent advances in near-optimal algorithms for the online resource allocation problem, which has been extensively studied, with diverse applications modeled through various linear program (LP) formulations. Representative examples include the secretary problem [15], online knapsack [4], network revenue management [17], network routing [6], and online matching [34]. Two primary input models are commonly considered in online LP research: i) the stochastic model, where each constraint column and objective coefficient is drawn independently from an unknown distribution, and ii) the random permutation model, where inputs arrive in a uniformly random order [35, 1, 27, 20]. Under a standard non-degeneracy assumption, logarithmic regret bounds have been established for problems such as quantity-based network revenue management [24, 23], general online LPs [29], and convex resource allocation [30]. More recent work has relaxed this assumption (e.g. [7, 40, 5, 25, 42, 3, 44]), leading to improved theoretical guarantees under broader and more realistic conditions. Our algorithm is new compared to the existing ones and we also relax the non-degeneracy assumption.

B Preliminary

We consider the bilinear saddle point problem defined as follows

$$\min_{\mathbf{x}\in\mathcal{X}=\Delta_{m_1}} \max_{\mathbf{y}\in\mathcal{Y}=\Delta_{m_2}} \mathbf{x}^\top A\mathbf{y}, \tag{8}$$

where $\Delta_n = \{\mathbf{x}\in\mathbb{R}^n : \sum_{i=1}^n \mathbf{x}_i = 1, \mathbf{x}_i \geq 0, \forall i\in[n]\}$ denotes the $(n-1)$-dimensional simplex and $A\in[-1,1]^{m_1\times m_2}$ is the payoff matrix. According to the celebrated minimax theorem [39], we know that $\min_{\mathbf{x}\in\mathcal{X}}\max_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top A\mathbf{y} = \max_{\mathbf{y}\in\mathcal{Y}}\min_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top A\mathbf{y}$. Define $\mathcal{X}^*\times\mathcal{Y}^*$ as the set of Nash Equilibria where $\mathcal{X}^* = \operatorname{argmin}_{\mathbf{x}\in\mathcal{X}}\max_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top A\mathbf{y}$ and $\mathcal{Y}^* = \operatorname{argmax}_{\mathbf{y}\in\mathcal{Y}}\min_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top A\mathbf{y}$. In the following, we introduce two definitions that measure of the closeness between a strategy (x,y) and the Nash Equilibria set $\mathcal{X}^*\times\mathcal{Y}^*$.

Definition 3 *We call a pair of strategy* $(\mathbf{x},\mathbf{y})\in\Delta_{m_1}\times\Delta_{m_2}$ *an* ϵ-close NE *if* $\operatorname{argmin}_{\mathbf{x}^*\in\mathcal{X}^*}\|\mathbf{x}-\mathbf{x}^*\|_2\leq\epsilon$ *and* $\operatorname{argmin}_{\mathbf{y}^*\in\mathcal{Y}}\|\mathbf{y}-\mathbf{y}^*\|_2\leq\epsilon$.

Definition 4 *Define the sub-optimality gap of $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ as* $\max_{\mathbf{y}' \in \Delta_{m_2}} \mathbf{x}^\top A \mathbf{y}' - \max_{\mathbf{y}' \in \Delta_{m_2}} x^{*\top} A \mathbf{y}'$ *and* $\min_{\mathbf{x}' \in \Delta_{m_1}} \mathbf{x}'^\top A \mathbf{y}^* - \min_{\mathbf{x}' \in \Delta_{m_1}} x'^\top A \mathbf{y}$ *(independent fo the choice of $x^* \in \mathcal{X}^*$ and $y^* \in \mathcal{Y}$). We call $x \in \mathcal{X}$ ($y \in \mathcal{Y}$) an ϵ-sub-optimality-gap NE if x's (y's) sub-optimality gap is no more than ϵ.*

Since we assume that $A \in [-1, 1]^{m_1 \times m_2}$, we know that an ϵ-close NE implies an ϵ-sub-optimality-gap NE, while the reverse is not true. In this paper, we consider the setting where only noisy bandit feedback is available. Specifically, the learner can only query an entry (i, j) of A and observe a sample $A_{i,j} + \eta \in [-1, 1]$ where η is a zero-mean noise.

Other Notations. For a vector $v \in \mathbb{R}^d$, let v_i denote its i-th entry, and for a subset $S \subseteq [d]$, let v_S denote the subvector of v containing the entries indexed by S. We say $v \succeq 0$ ($v \preceq 0$) if $v_i \geq 0$ ($v_i \leq 0$) for all $i \in [d]$. Denote $\mathbf{e}^n$ as the all-ones vector in n-dimensional space and $\mathbf{0}$ as the all-zeros vector in an appropriate dimension. For a matrix $M \in \mathbb{R}^{n_1 \times n_2}$ and two sets $S_1 \subseteq [n_1]$, $S_2 \subseteq [n_2]$, let $M_{S_1,:} \in \mathbb{R}^{|S_1| \times n_2}$ denote the submatrix containing the rows of M indexed by S_1. Similarly, let $M_{:,S_2} \in \mathbb{R}^{n_1 \times |S_2|}$ denote the submatrix containing the columns of M indexed by S_2, and let $M_{S_1,S_2} \in \mathbb{R}^{|S_1| \times |S_2|}$ denote the one containing the entries with the row indices in S_1 and the column indices in S_2. We also let $m = m_1 m_2$ for notational convenience.

B.1 Reformulation as Linear Programming

In this section, we present a reformulation of the bilinear saddle-point problem defined in Equation (8). We show that in order to solve Equation (8), we can focus on two linear programs, which can be regarded as the primal-dual form to each other. We then exploit the specific structures of the linear programming reformulations to derive our algorithms. Specifically, for a fixed $\mathbf{x}$, the inner maximization problem over $\mathbf{y}$ can be viewed as a linear program, the dual of which can be written as

$$\min_{\mu \in \mathbb{R}} \mu \quad \text{s.t. } \mu \cdot \mathbf{e}^{m_2} \geq A^\top \mathbf{x}. \tag{9}$$

Further minimizing over $\mathbf{x} \in \mathcal{X}$ reaches the following "primal" linear program.

$$V^{\text{Prime}} = \min_{\mathbf{x} \succeq 0, \mu \in \mathbb{R}} \mu \quad \text{s.t. } \mu \cdot \mathbf{e}^{m_2} \succeq A^\top \mathbf{x}, \ (\mathbf{e}^{m_1})^\top \mathbf{x} = 1. \tag{10}$$

According to the celebrated minimax theorem, Equation (8) is equivalent to $\max_{\mathbf{y} \in \mathcal{Y}} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top A \mathbf{y}$. Then, following a similar process, we obtain the "dual" linear programming that computes $\mathbf{y}$.

$$V^{\text{Dual}} = \max_{\nu \geq 0, \mathbf{y} \succeq 0} \nu \quad \text{s.t. } \nu \cdot \mathbf{e}^{m_1} \preceq A\mathbf{y}, \ (\mathbf{e}^{m_2})^\top \mathbf{y} = 1. \tag{11}$$

Direct calculation shows that LP in Equation (10) and LP in Equation (11) are primal-dual to each other. Moreover, in order to solve Equation (8), classic LP analysis (e.g. Theorem 16.5 in [18]) shows that it suffices to find an optimal solution $\mathbf{x}^*$ to LP Equation (10) and an optimal solution $\mathbf{y}^*$ to LP Equation (11), such that $\mathbf{x}^*$ and $\mathbf{y}^*$ are the corresponding primal-dual solutions.

Lemma 1 *For any optimal solution to the primal LP Equation* (10)*, denoted by* $(\mathbf{x}^*, \mu^*)$*, and the corresponding optimal dual solution to LP Equation* (11)*, denoted by* $(\mathbf{y}^*, \nu^*)$*,* $(\mathbf{x}^*, \mathbf{y}^*)$ *is an optimal solution to Equation* (8)*.*

In the following, we first present in Section C a method for identifying the support of an NE, denoted as $\mathcal{I}^* \subseteq \mathcal{X}^*$ and $\mathcal{J}^* \subseteq \mathcal{Y}^*$, using $\mathcal{O}(\frac{m}{\delta^2})$ samples, where δ is a problem-dependent constant defined in Section C. Then, in Section D, we demonstrate how to compute the NE supported on $\mathcal{I}^* \times \mathcal{J}^*$.

C Saddle Point Support Identification

In this section, we develop methods to identify the *support* for an optimal solution $\mathbf{x}^*$ to the primal LP Equation (10). Following standard LP theory, we know that when solving an LP, we can restrict to the corner points of the feasible region, which is also referred to as *basic solution*. The following lemma shows that there always exists two index sets $\mathcal{I}^*$ and $\mathcal{J}^*$ such that one optimal solution to the LP V^{Prime} can be fully characterized by the sets $\mathcal{I}^*$ and $\mathcal{J}^*$ as the solution to a set of linear equations.

Theorem 4 *There exists index sets $\mathcal{I}^* \subset [m_1]$ and index sets $\mathcal{J}^* \subset [m_2]$, such that $|\mathcal{I}^*| = |\mathcal{J}^*| = d$, for some integer $d \le \min\{m_1, m_2\}$. Also, for the given sets $\mathcal{I}^*$ and $\mathcal{J}^*$, there exists an optimal solution $(\mathbf{x}^*, \mu^*)$ to Equation (10) that satisfy*

$$A^\top_{\mathcal{I}^*,\mathcal{J}^*} \cdot \mathbf{x}^*_{\mathcal{I}^*} = \mu^* \cdot \mathbf{e}^{|\mathcal{J}^*|} \ \ \textit{and} \ \ (\mathbf{e}^{m_1})^\top \mathbf{x}^* = 1, \tag{12}$$

and $\mathbf{x}^*_{\mathcal{I}^{*c}} = 0$, where $\mathcal{I}^{*c} = [m_1] \setminus \mathcal{I}^*$, and $\mathcal{J}^{*c} = [m_2] \setminus \mathcal{J}^*$ denote the complementary sets.

In order to identify the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$ from noisy observations, we first sample each entry $A_{i,j}$ a number of times and construct an empirical game matrix $\widehat{A}$ with each entry the empirical average of the N observations, and construct the primal LP of Equation (10) using $\widehat{A}$. Next, we successively shrink the size of x-player's support based on the objective value $\widehat{V}^{\text{Prime}}_{\mathcal{I}}$ defined in Equation (13). Specifically, $\widehat{V}^{\text{Prime}}_{\mathcal{I}}$ is defined as the empirical primal LP with non-zero values only on $\mathcal{I}$ and we initialize the support set to be $\mathcal{I} = [m_1]$. At each time, we sequentially pick each $i \in \mathcal{I}$ and compare the objective value of $\widehat{V}^{\text{Prime}}_{\mathcal{I}}$ with $\widehat{V}^{\text{Prime}}_{\mathcal{I}\setminus\{i\}}$. If these two quantities are the same, this means that there exists an NE that is not supported on i and we set $\mathcal{I} \leftarrow \mathcal{I}\setminus\{i\}$; otherwise, we keep i in $\mathcal{I}$. We then obtain the support set $\mathcal{I}^*$ after iterating over all $i \in [m_1]$.

$$\begin{aligned} V^{\text{Prime}}_{\mathcal{I}} = \min \quad & \mu \\ \text{s.t.} \quad & \mu \cdot \mathbf{e}^{m_2} \succeq A^\top \mathbf{x} \\ & (\mathbf{e}^{m_1})^\top \mathbf{x} = 1 \\ & \mathbf{x}_{\mathcal{I}^c} = \mathbf{0} \\ & \mathbf{x} \succeq \mathbf{0}, \mu \in \mathbb{R}, \end{aligned} \qquad \begin{aligned} \widehat{V}^{\text{Prime}}_{\mathcal{I}} = \quad \min \quad & \mu \\ \text{s.t.} \quad & \mu \cdot \mathbf{e}^{m_2} \succeq \widehat{A}^\top \mathbf{x} \\ & (\mathbf{e}^{m_1})^\top \mathbf{x} = 1 \\ & \mathbf{x}_{\mathcal{I}^c} = \mathbf{0} \\ & \mathbf{x} \succeq \mathbf{0}, \mu \in \mathbb{R}. \end{aligned} \tag{13}$$

After identifying the set $\mathcal{I}^*$, we also need to find the set $\mathcal{J}^*$, which is equivalent to see what is the support for the dual LP. We apply a similar approach to find $\mathcal{J}^*$. Specifically, based on $\mathcal{I}^*$, we define the dual LP $\widehat{V}^{\text{Dual}}_{\mathcal{I}^*,\mathcal{J}}$ in Equation (14) in which we only include the rows of $\widehat{A}$ whose indices are in $\mathcal{I}^*$ and set the dual variables outside $\mathcal{J}$ to be 0. We initialize $\mathcal{J} = [m_2]$ and successively eliminate an element $j \in \mathcal{J}$ if $\widehat{V}_{\mathcal{I}^*,\mathcal{J}\setminus\{j\}}$ equals to $\widehat{V}^{\text{Prime}}_{\mathcal{I}^*,\mathcal{J}}$. Finally, we obtain $\mathcal{J}^*$ after iterating over all $j \in [m_2]$. The pseudo code of the algorithm is shown in Algorithm 2. Note that it is possible that there are multiple optimal basis of $\mathcal{I}^*$ and $\mathcal{J}^*$ and we show later that our approach essentially identifies one particular optimal basis, since our goal is to find one NE of the game. In the following, we provide both instance-dependent and worst-case sample complexity guarantees.

$$\begin{aligned} V^{\text{Dual}}_{\mathcal{I}^*,\mathcal{J}} = \max \quad & \nu \\ \text{s.t.} \quad & \nu \cdot \mathbf{e}^{|\mathcal{I}^*|} \preceq A_{\mathcal{I}^*,:} \cdot \mathbf{y} \\ & (\mathbf{e}^{m_2})^\top \mathbf{y} = 1 \\ & \mathbf{y}_{\mathcal{J}^c} = \mathbf{0} \\ & \mathbf{y} \succeq 0, \nu \in \mathbb{R}, \end{aligned} \qquad \begin{aligned} \widehat{V}^{\text{Dual}}_{\mathcal{I}^*,\mathcal{J}} = \quad \max \quad & \nu \\ \text{s.t.} \quad & \nu \cdot \mathbf{e}^{|\mathcal{I}^*|} \preceq \widehat{A}_{\mathcal{I}^*,:} \cdot \mathbf{y} \\ & (\mathbf{e}^{m_2})^\top \mathbf{y} = 1 \\ & \mathbf{y}_{\mathcal{J}^c} = \mathbf{0} \\ & \mathbf{y} \succeq 0, \nu \in \mathbb{R}. \end{aligned} \tag{14}$$

C.1 Instance-Dependent Sample Complexity Guarantee

In this section, we show that Algorithm 2 can successfully detect the optimal support index sets $\mathcal{I}^*$ and $\mathcal{J}^*$ with a high probability with the number of samples a constant with respect to certain instance-dependent quantity. Specifically, we define $\delta > 0$ as follows

Definition 5 *Define $\delta_1 \triangleq \min_{\mathcal{I}}\{V^{\text{Prime}} - V^{\text{Prime}}_{\mathcal{I}} : V^{\text{Prime}} - V^{\text{Prime}}_{\mathcal{I}} > 0\}$ to be the minimum non-zero primal gap where V^{Prime} is defined in Equation (10) and $V^{\text{Prime}}_{\mathcal{I}}$ is defined in Equation (13). Define $\delta_2 \triangleq \min_{\mathcal{I},\mathcal{J}}\{V^{\text{Dual}}_{\mathcal{I},\mathcal{J}} - V^{\text{Dual}}_{\mathcal{I},[m_2]} : V^{\text{Dual}}_{\mathcal{I},\mathcal{J}} - V^{\text{Dual}}_{\mathcal{I},[m_2]} > 0\}$ to be the minimum non-zero dual gap where $V^{\text{Dual}}_{\mathcal{I},\mathcal{J}}$ is defined in Equation (14). Define $\delta = \min\{\delta_1, \delta_2\}$.*

Specifically, δ measures the minimum non-zero gap to the optimal objective values (for both the primal and the dual) when the value on a support of the NE is set to be 0.

Algorithm 2 Support Identification

Input: the empirical matrix $\widehat{A}$ with each entry the average of $\frac{N}{m}$ noisy observations
Initialize $\mathcal{I} = [m_1]$ and $\mathcal{J} = [m_2]$.
Compute the objective value of $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$ as in Equation (13). Set this value to be V.
for $i \in \mathcal{I}$ **do**
    Let $\mathcal{I}' = \mathcal{I} \backslash \{i\}$ and compute the value of $\widehat{V}_{\mathcal{I}'}^{\mathrm{Prime}}$.
    if $V = \widehat{V}_{\mathcal{I}'}^{\mathrm{Prime}}$ **then** $\mathcal{I} = \mathcal{I}'$.
for $j \in \mathcal{J}$ **do**
    Let $\mathcal{J}' = \mathcal{J} \backslash \{j\}$ and compute the value of $\widehat{V}_{\mathcal{I},\mathcal{J}'}^{\mathrm{Dual}}$.
    Compute the smallest singular value of the matrix $\begin{bmatrix} \widehat{A}_{\mathcal{I},\mathcal{J}'} & -\mathbf{e}^{|\mathcal{I}|} \\ (\mathbf{e}^{|\mathcal{J}'|})^\top & 0 \end{bmatrix}$, denoted as $\widehat{\sigma}_{\mathcal{I},\mathcal{J}'}$.
    if $V = \widehat{V}_{\mathcal{I},\mathcal{J}'}^{\mathrm{Dual}}$ and $\widehat{\sigma}_{\mathcal{I},\mathcal{J}'} > 2|\mathcal{I}| \cdot \mathrm{Rad}(N'/m, \varepsilon/m)$ with $\mathrm{Rad}(N'/m, \varepsilon/m) = \sqrt{\frac{m \cdot \log(2m/\varepsilon)}{2N'}}$. **then** $\mathcal{J} = \mathcal{J}'$.
Output: the sets of indices $\mathcal{I}$ and $\mathcal{J}$.

The next theorem shows that when the number of samples N for each entry is at least $\Omega(\frac{1}{\delta^2})$, then with a high probability, $\mathcal{I}$ and $\mathcal{J}$ output by Algorithm 2 satisfies Theorem 4.

Theorem 5 *For any $\varepsilon > 0$, as long as $N \geq \frac{2m\log(2m/\epsilon)}{\delta^2}$, with probability at least $1-\varepsilon$, the outputs $\mathcal{I}$ and $\mathcal{J}$ of Algorithm 2 satisfy the conditions described in Theorem 4.*

The proof is deferred to Section G. To provide a proof sketch, following standard concentration inequalities, with $N \geq \frac{2m\log(2m/\varepsilon)}{\delta^2}$, each entry of the game matrix A is approximated with an error of $\frac{\delta}{2}$ with probability at least $1-\varepsilon$. Under this high probability event, we can show that the gap between the optimal objective value of $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$ and $V_{\mathcal{I}}^{\mathrm{Prime}}$, as well as the dual $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ and $V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ for all $\mathcal{I} \subseteq [m_1]$ and $\mathcal{J} \subseteq [m_2]$, is smaller than $\frac{\delta}{2}$. Then, we are able to show that the optimal basis of the empirical LP $\widehat{V}^{\mathrm{Prime}}$ estimated from samples is also an optimal basis of the original LP V^{Prime}.

Two remarks are as follows. First, the proof sketch above shows that we can also directly compute an optimal basis of $\widehat{V}^{\mathrm{Prime}}$ and, when $N \geq \frac{2m\log(2m/\epsilon)}{\delta^2}$, this basis is also optimal to the original LP V^{Prime} with probability $1-\varepsilon$. While there are standard methods to compute the optimal basis of a given LP (such as the simplex method), these can require exponential time in the worst case. In contrast, the approach in Algorithm 2, which computes the LP value to select the basis, achieves a polynomial time complexity since linear programming itself can be solved in polynomial time. Second, while we require the knowledge of δ to decide N in Theorem 5, we show in Section L that we can also estimate δ within a factor of 2 using $\mathcal{O}(\frac{m\log(m/\epsilon)}{\delta^2})$ samples *without knowing the parameter* δ, meaning that we can find the support set without the knowledge of δ.

C.2 δ-Independent Sample Complexity Guarantee

In Section C.1, we show that with $N = \Theta(\frac{m\log(m/\epsilon)}{\delta^2})$ samples, Algorithm 2 identifies the desired optimal support $\mathcal{I}^*$ and $\mathcal{J}^*$ with a probability at least $1-\varepsilon$. However, when δ is small, to obtain $\mathcal{I}^*$ and $\mathcal{J}^*$ that satisfies Theorem 4, the number of samples we need can be very large. Therefore, in this section, we aim to provide a δ-independent sample complexity guarantee on the obtained support $\mathcal{I}$ and $\mathcal{J}$ output by Algorithm 2. In this case, instead of proving that Algorithm 2 finds the optimal support with high probability, we show that with $N = \mathcal{O}(m/\epsilon^2)$ samples, Algorithm 2 finds $\mathcal{I}$ and $\mathcal{J}$ such that the optimal solution supported on $\mathcal{I}$ has a sub-optimality gap bounded at most ε. To this end, for possible basis sets $\mathcal{I} \subset [m_1]$ and $\mathcal{J} \subset [m_2]$ such that $|\mathcal{I}| = |\mathcal{J}|$, we define $\mathbf{x}^*(\mathcal{I},\mathcal{J})$ as the solution such that $\mathbf{x}^*_{\mathcal{I}^c}(\mathcal{I},\mathcal{J}) = 0$ and the non-zero elements $\mathbf{x}^*_{\mathcal{I}}(\mathcal{I},\mathcal{J})$ are the solution to the following equation:

$$\begin{bmatrix} A_{\mathcal{I},\mathcal{J}}^\top & -\mathbf{e}^{|\mathcal{J}|} \\ (\mathbf{e}^{|\mathcal{I}|})^\top & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}^*_{\mathcal{I}}(\mathcal{I},\mathcal{J}) \\ \mu^*(\mathcal{I},\mathcal{J}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{15}$$

The following theorem proves the sub-optimality gap of $\mathbf{x}^*(\mathcal{I}^N, \mathcal{J}^N)$, where $\mathcal{I}^N$ and $\mathcal{J}^N$ denote the output of Algorithm 2 with sample size N.

Theorem 6 *Let $\mathrm{Rad}(n, \epsilon) = \mathcal{O}\left(\sqrt{\frac{\log(1/\epsilon)}{n}}\right)$. For a fixed sample size N, denote by $\mathcal{I}^N$ and $\mathcal{J}^N$ the output of Algorithm 2. Then, for any $\varepsilon > 0$, with a probability at least $1 - \varepsilon$, we have that*

$$|V^{\mathrm{Prime}} - V^{\mathrm{Prime}}_{\mathcal{I}^N}| \le \mathrm{Rad}(N/m, \varepsilon/m) \ \text{ and } \ |V^{\mathrm{Prime}} - V^{\mathrm{Dual}}_{\mathcal{I}^N, \mathcal{J}^N}| \le \mathrm{Rad}(N/m, \varepsilon/m). \tag{16}$$

Moreover, with probability at least $1 - \varepsilon$, the solution $\mathbf{x}^*(\mathcal{I}^N, \mathcal{J}^N)$ forms a feasible solution to V^{Prime} with a sub-optimality gap bounded by $\mathcal{O}\left(\widehat{\kappa} \cdot \mathrm{Rad}(N/m, \varepsilon/m)\right)$, where $\widehat{\kappa} > 0$ is the condition number of the matrix $\begin{bmatrix} \widehat{A}^{\top}_{\mathcal{I}^N, \mathcal{J}^N} & -\mathbf{e}^{|\mathcal{J}^N|} \\ (\mathbf{e}^{|\mathcal{I}^N|})^{\top} & 0 \end{bmatrix}$.

The proof is deferred to Section H and we provide a proof sketch here. Specifically, similar to the proof sketch for Theorem 5, based on the high probability event that each entry of A is approximated within an error of $\mathrm{Rad}(N/m, \varepsilon/m)$, we know that the objective value gap between V^{Prime} and $\widehat{V}^{\mathrm{Prime}}$, as well as $V^{\mathrm{Prime}}_{\mathcal{I}^N}$ and $\widehat{V}^{\mathrm{Prime}}_{\mathcal{I}^N}$, are bounded by $\mathrm{Rad}(N/m, \varepsilon/m)$. Since Algorithm 2 guarantees that $\widehat{V}^{\mathrm{Prime}}_{\mathcal{I}^N} = \widehat{V}^{\mathrm{Prime}}$, we know that the gap between $V^{\mathrm{Prime}}_{\mathcal{I}^N}$ and V^{Prime} is bounded by $\mathrm{Rad}(N/m, \epsilon/m)$, proving the first half of Equation (16). The second half can be obtained via a similar argument. To prove the sub-optimality gap, we control the distance between $x^*(\mathcal{I}^N, \mathcal{J}^N)$ and $\widehat{x}^*(\mathcal{I}^N, \mathcal{J}^N)$, which is (part of) the solution of $\begin{bmatrix} \widehat{A}^{\top}_{\mathcal{I}^N, \mathcal{J}^N} & -\mathbf{e}^{|\mathcal{J}^N|} \\ (\mathbf{e}^{|\mathcal{I}^N|})^{\top} & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$. Following a standard perturbation analysis of the linear program, we can show that $\|\widehat{x}^*(\mathcal{I}^N, \mathcal{J}^N) - x^*(\mathcal{I}^N, \mathcal{J}^N)\|_2 \le \mathcal{O}\left(\widehat{\kappa} \cdot \mathrm{Rad}(N/m, \epsilon/m)\right)$. The remaining analysis follows a direct derivation.

D LP-Resolving Based Algorithm

In the previous section, we described how to identify one optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$. In this section, we describe how to approximate the optimal solutions $\mathbf{x}^*$ and $\mathbf{y}^*$ that corresponds to the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$. Note that the optimal solution $\mathbf{x}^*$ enjoys the following structure: $\mathbf{x}^*_{\mathcal{I}^{*c}} = 0$ and the other elements $\mathbf{x}^*_{\mathcal{I}^*}$ can be given as the solution to

$$\begin{bmatrix} A^{\top}_{\mathcal{I}^*, \mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^{\top} & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}^*_{\mathcal{I}^*} \\ \mu^* \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{17}$$

While Algorithm 1 can determine the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$, since matrix A is unknown, and we can only access noisy observations, we must construct estimates of A using sequentially collected samples to approximately solve Equation (17).

Our approach is formalized in Algorithm 3. Specifically, in each iteration n, we observe a noisy matrix entry $\tilde{A}_{n,i_n,j_n}$ with $\mathbb{E}[\tilde{A}_{n,i_n,j_n}] = A_{i_n,j_n}$ and (i_n, j_n) uniformly drawn from $\mathcal{I}^* \times \mathcal{J}^*$. A naive approach would be to directly use these samples to form an empirical estimate of $A_{\mathcal{I}^*, \mathcal{J}^*}$ and compute the optimal basis. However, this would result in estimation error of $\Theta(1/\sqrt{N})$ by standard concentration bounds, leading to a suboptimal $\Theta(1/\sqrt{N})$ approximate Nash Equilibrium.

To achieve a better convergence rate, Algorithm 3 is inspired by the LP resolving algorithms for online resource allocation problems [40**?** , 25, 30] that can achieve $\mathcal{O}(\log N)$ regret after N iterations. Specifically, we interpret Equation (17) as a resource allocation problem: the right-hand-side represents resources, variables $\mathbf{x}$ represent actions, and the left-hand-side matrix represents resource consumption. The goal is to determine actions $\{\mathbf{x}_n\}_{n \in [N]}$ such that resources are appropriately consumed over N iterations.

The crucial point of Algorithm 2's effectiveness lies in its adaptive resource management mechanism (Line 8 of Algorithm 3). By dynamically updating $\mathbf{a}^n$ according to Equation (19), we create a self-correcting system that forms and solves the LP for $\tilde{\mathbf{x}}^n$ as specified in Equation (18). This resolving procedure ensures that $\mathbf{a}^n$ converges to $\mathbf{a}$ with a gap bounded by $O(1/(N - n + 1))$ – a critical factor in achieving our instance-dependent guarantee. For example, when a binding constraint $j \in \mathcal{J}^*$

Algorithm 3 The Resolving Algorithm

Input: the number of samples N, the support $\mathcal{I}^*$ and $\mathcal{J}^*$ with $|\mathcal{I}^*| = |\mathcal{J}^*|$, $L = 4$.
Initialize $\mathcal{H}^1 = \emptyset$, $\mathbf{a}^1 = \mathbf{0} \in \mathbb{R}^{|\mathcal{J}^*|}$
for $n = 1, \ldots, N$ **do**
  Construct estimates $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}$ using the dataset $\mathcal{H}^n$.
  Construct a solution $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ such that $\tilde{\mathbf{x}}^n_{\mathcal{I}^{*c}} = 0$ and $\tilde{\mathbf{x}}^n_{\mathcal{I}^*}$ is the solution to

$$\begin{bmatrix} \widehat{A}^\top_{\mathcal{I}^*,\mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_{\mathcal{I}^*} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{a}^n/(N-n+1) \\ 1 \end{bmatrix}. \tag{18}$$

  Project $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ to the set $\{(\mathbf{x}, \mu) : \mathbf{x} \geq 0, \|(\mathbf{x}, \mu)\|_1 \leq L\}$ to obtain $(\mathbf{x}^n, \mu^n)$.
  Observe $\tilde{A}_{n,i_n,j_n} = A_{i_n,j_n} + \eta_n$, where i_n and j_n are uniformly sampled from $\mathcal{I}^*$ and $\mathcal{J}^*$.
  Update $\mathcal{H}^{n+1} = \mathcal{H}^n \cup \{(i_n, j_n, \tilde{A}_{n,i_n,j_n})\}$ and compute $\mathbf{a}^{n+1}$ as

$$\mathbf{a}^{n+1} = \mathbf{a}^n - |\mathcal{J}^*| \cdot |\mathcal{I}^*| \cdot \tilde{A}_{i_n,j_n} \cdot x^n_{i_n} \cdot \mathbf{h}_{j_n} + \mu^n \cdot \mathbf{e}^{|\mathcal{J}^*|}, \tag{19}$$

  where $\mathbf{h}_{j_n} \in \mathbb{R}^{|\mathcal{J}^*|}$ denotes a vector with j_n-th position being 1 and other positions being 0.
Compute $\bar{\mathbf{x}} = \frac{1}{N} \cdot \sum_{n=1}^N \mathbf{x}^n$ and output $\bar{\mathbf{x}}$.

has a value below its target $\mathbf{a}_j$ under the current action $\sum_{n'=1}^n \mathbf{x}^{n'}/n$, the algorithm compensates by setting $\mathbf{a}^n_j/(N-n+1)$ greater than $\mathbf{a}_j$, thereby increasing the constraint value for $\tilde{\mathbf{x}}^{n+1}$. Through this continuous feedback loop, the algorithm dynamically adjusts for constraint violations, enabling the significantly faster convergence rate of $O(1/(N-n+1))$ compared to the naive approach.

In the following, similar to Section C, we first introduce our instance-dependent sample complexity for Algorithm 3 in Section D.1, followed by a δ-independent sample complexity in Section D.2.

D.1 Instance-dependent Sample Complexity

We now show that our Algorithm 3 achieves an instance-dependent sample complexity for approximating the optimal solution to the primal LP problem V^{Prime}. A key step in our analysis is to characterize how $\mathbf{a}^n$ behaves during the execution of Algorithm 3. For notational convenience, we define

$$\tilde{\mathbf{a}}^n = \frac{\mathbf{a}^n}{N-n+1}. \tag{20}$$

The key is to show that the stochastic process of $\tilde{\mathbf{a}}^n$ enjoys nice concentration properties such that it will stay within a small neighborhood of their initial value $\mathbf{0}$ for a sufficiently long time. We denote by τ the first time index at which $\|\tilde{\mathbf{a}}^n\|_\infty$ exceeds certain problem-dependent constant (formally defined in Equation (63)). Then, we can show that the gap $\mathrm{argmin}_{\mathbf{x}^* \in \mathcal{X}^*} \|\mathbb{E}[\bar{\mathbf{x}}] - \mathbf{x}^*\|_2$ can be upper bounded by $\frac{\mathbb{E}[N-\tau]}{N}$. Thus, it only remains to upper bound $\mathbb{E}[N-\tau]$. From the update rule in Equation (19), we have

$$\tilde{\mathbf{a}}^{n+1} = \tilde{\mathbf{a}}^n + \frac{\tilde{\mathbf{a}}^n - |\mathcal{J}^*| \cdot |\mathcal{I}^*| \cdot \tilde{A}_{i_n,j_n} \cdot x^n_{i_n} \cdot \mathbf{h}_{j_n} + \mu^n \cdot \mathbf{e}^{|\mathcal{J}^*|}}{N-n}. \tag{21}$$

Note that since i_n (resp. j_n) is sampled from a uniform distribution over $\mathcal{I}^*$ (resp. $\mathcal{J}^*$) , we have that

$$\mathbb{E}_{\eta_n,i_n,j_n}\left[|\mathcal{J}^*| \cdot |\mathcal{I}^*| \cdot \tilde{A}_{i_n,j_n} \cdot x^n_{i_n} \cdot \mathbf{h}_{j_n}\right] = A^\top_{\mathcal{I}^*,\mathcal{J}^*} \cdot \mathbf{x}^n_{\mathcal{I}^*}. \tag{22}$$

Note that we can show it is automatically satisfied $\|(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)\|_1 \leq L$ when $n \geq N_0'$ for some fixed $N_0' = O(\log N)$, and thus $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n) = (\mathbf{x}^n, \mu^n)$. Then we know when $n \geq N_0'$, $x^n_{\mathcal{I}^*}$ is the solution to Equation (18) and we know that $\widehat{A}^\top_{\mathcal{I}^*,\mathcal{J}^*} x^n_{\mathcal{I}^*} - \mu^n \cdot \mathbf{e}^{|\mathcal{J}^*|} = \tilde{\mathbf{a}}^n$. Therefore, if $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*} = A_{\mathcal{I}^*,\mathcal{J}^*}$, then $\tilde{\mathbf{a}}^{n+1}$ will have the same expectation as $\tilde{\mathbf{a}}^n$ such that they become a martingale. However, this is not exactly true since we have estimation error over $A_{\mathcal{I}^*,\mathcal{J}^*}$, and we only use their estimates to compute $\mathbf{x}^n$. Nevertheless, we can show that $\tilde{\mathbf{a}}^n$ behaves as sub-martingales with estimation error over $A_{\mathcal{I}^*,\mathcal{J}^*}$ decreases. Then, from the concentration property of the sub-martingale, we prove that $\tilde{\mathbf{a}}^n$ concentrates around $\mathbf{0}$. In this way, we prove that $\mathbb{E}[N-\tau]$ can be upper bounded by $O(\log N)$. The following theorem shows our formal theoretical bound by applying Algorithm 2 and Algorithm 3.

Theorem 7 *For any $\epsilon > 0$, with a sample complexity of $N = \mathcal{O}\left(\left(\frac{d^{5/2}}{\sigma^3} + \frac{m}{\delta^2}\right) \cdot \frac{\log(m/\varepsilon)}{\varepsilon}\right)$, where $\delta = \min\{\delta_1, \delta_2\}$ with δ_1 and δ_2 given in Definition 5 and $d = |\mathcal{I}^*| = |\mathcal{J}^*|$, applying Algorithm 2 and Algorithm 3 guarantees that $\operatorname{argmin}_{\mathbf{x}^* \in \mathcal{X}^*} \|\mathbb{E}[\bar{\mathbf{x}}] - \mathbf{x}^*\|_2 \leq \varepsilon$. Here, $\sigma > 0$ is the minimum singular value of the full-rank matrix A^* defined as*

$$A^* = \begin{bmatrix} A^\top_{\mathcal{I}^*, \mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top, & 0 \end{bmatrix}. \tag{23}$$

Note that the sample complexity presented in Theorem 7 again depends on the gap δ. However, neither our Algorithm 2 nor Algorithm 3 requires the knowledge of δ since we can estimate δ within a factor of 2 according to Theorem 10. The bound in Theorem 7 also depends on the parameter σ, which is the smallest singular values of the matrix A^* in Equation (23). Note that in previous literature that develops the LP resolving algorithm with instance-dependent guarantees (e.g. [40, 28, 25]), the conditional number of the constraint matrix will also show up in the final bounds. Therefore, we regard the existence of σ as the consequence of adopting the resolving algorithms. In addition, similar to the case for δ, we show in Theorem 11 in Section L that we can also estimate σ within a factor of 2 using observed samples, meaning that our algorithm also *does not* require the knowledge of σ.

D.2 δ-Independent Sample Complexity

In this section, we show that Algorithm 3 also achieves an δ-independent sub-optimality gap with N samples, complementing the instance-dependent sample complexity proven in Section D.2. Specifically, we have the following theorem:

Theorem 8 *For a fixed sample size N, denote by $\bar{\mathbf{x}}$ the output of Algorithm 3 (after processing Algorithm 2 with the same amount of data). We know that $\mathbb{E}[\bar{\mathbf{x}}]$ is a feasible solution to V^{Prime} with an suboptimality gap upper bounded by*

$$\mathcal{O}\left(\frac{d^2}{\sigma} \cdot \left(1 + \frac{d}{\sigma^2 \cdot L^2}\right) \cdot \frac{\log(N)}{N} + \widehat{\kappa} \cdot \sqrt{\frac{dm \cdot \log(N)}{N}}\right),$$

where $\widehat{\kappa}$ is defined to be the same as Theorem 6, $d = |\mathcal{I}^*| = |\mathcal{J}^*|$, and σ is the minimum singular value of A^* defined in Theorem 7.

We provide a proof sketch of Theorem 8 with the full proof deferred to Section J. Note that following Theorem 6, for a fixed sample size N, Algorithm 2 outputs index sets $\hat{\mathcal{I}}, \hat{\mathcal{J}}$ such that the variable $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$, as defined in Equation (15), forms a feasible solution to V^{Prime} with an suboptimality gap $\widetilde{\mathcal{O}}(1/\sqrt{N})$. Then, we use the index sets $\hat{\mathcal{I}}, \hat{\mathcal{J}}$ as the input to Algorithm 3. Note that given the index sets $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$, the resolving steps in Algorithm 3 essentially approximate the solution to the linear equations given in Equation (15). Following the same procedure as the proof of Theorem 7, we can show that after N steps of resolving, the gap between $\mathbb{E}[\bar{\mathbf{x}}]$ and $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$ is upper bounded by $\widetilde{\mathcal{O}}(1/N)$, corresponding to the first term in Theorem 8. Therefore, we know that $\mathbb{E}[\bar{\mathbf{x}}]$ forms a feasible solution to V^{Prime} with an suboptimality gap bounded by $\widetilde{\mathcal{O}}(1/\sqrt{N})$.

E Combining Two Steps Together

Now we analyze the overall performance of our algorithm. Specifically, combining Theorem 5 and Theorem 6 for the support identification step and Theorem 7 and Theorem 8 for the LP resolving step, we show in the following theorem that given $2N$ samples as long as $N = \widetilde{\mathcal{O}}\left(\min\left\{\frac{m}{\delta^2 \cdot \varepsilon}, \frac{m}{\varepsilon^2}\right\}\right)$, Algorithm 2 followed by Algorithm 3 is able to find an ε-suboptimality-gap NE.

Theorem 9 *For a fixed accuracy level $\varepsilon > 0$, denote by $\bar{\mathbf{x}}$ the output of Algorithm 3 with N samples (after processing Algorithm 2 with another N samples). We know that $\mathbb{E}[\bar{\mathbf{x}}]$ forms a feasible solution to V^{Prime}, with an suboptimality gap upper bounded by ε, as long as the sample complexity bound holds*

$$N = \mathcal{O}\left(\frac{d^3}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + m \cdot \min\left\{\frac{d \cdot \widehat{\kappa}^2}{\varepsilon^2}, \frac{1}{\delta^2 \cdot \varepsilon}\right\} \cdot \log(m/\varepsilon)\right),$$

where $\delta = \min\{\delta_1, \delta_2\}$ with δ_1 and δ_2 defined in Definition 5, and σ the minimum singular value of A^* defined in Theorem 7, and $\widehat{\kappa}$ given in the statement of Theorem 6.

The full proof is deferred to Section K. We remark again that our algorithm does not require the knowledge of σ and δ since both parameters can be well approximated using the same sample complexity as shown in Section L.

E.1 Results for the Dual Player

In this section, we briefly discuss how we derive an approximate NE for the column player. Specifically, we apply the exact same procedure, Algorithm 2 followed by Algorithm 3, but instead to the dual LP V^{Dual} in Equation (11). We then obtain the output of Algorithm 3 as $\bar{\mathbf{y}}$ to approximate an optimal solution $\mathbf{y}^*$ for the dual player. Note that all our previous results for the primal LP in Equation (10) still applies to the dual LP. The only difference is that the problem-dependent constants are now defined for the dual LP instead. For example, the problem parameters presented in Definition 5 can be modified into the following:

Definition 6 *Define $\delta_1' \triangleq \min_{\mathcal{I}}\{V^{\mathrm{Dual}} - V_{\mathcal{I}}^{\mathrm{Dual}} : V^{\mathrm{Dual}} - V_{\mathcal{I}}^{\mathrm{Dual}} > 0\}$ to be the minimum non-zero primal gap where V^{Dual} is defined in Equation* (11) *and we further restrict $\mathbf{y}_{\mathcal{I}} = \mathbf{0}$ to obtain $V_{\mathcal{I}}^{\mathrm{Dual}}$. Define $\delta_2' \triangleq \min_{\mathcal{I},\mathcal{J}}\{V_{\mathcal{I},\mathcal{J}}^{\mathrm{Prime}} - V_{\mathcal{I},[m_1]}^{\mathrm{Prime}} : V_{\mathcal{I},\mathcal{J}}^{\mathrm{Prime}} - V_{\mathcal{I},[m_1]}^{\mathrm{Prime}} > 0\}$ to be the minimum non-zero dual gap. Define $\delta' \triangleq \min\{\delta_1', \delta_2'\}$.*

Similarly, we are able to define σ', d', and $\widehat{\kappa}'$ correspondingly to the dual system. Then, following the same procedure as Theorem 7, we know that with a sample complexity bound of $N' = \mathcal{O}\left(\left(\frac{d'^{5/2}}{\sigma'^3} + \frac{m}{\delta'^2}\right) \cdot \frac{\log(m/\epsilon)}{\epsilon}\right)$, Algorithm 3 outputs a solution $\bar{\mathbf{y}}$ such that $\|\mathbf{y}^* - \mathbb{E}[\bar{\mathbf{y}}]\|_2 \le \varepsilon$, for an optimal dual solution $\mathbf{y}^*$. Therefore, applying Algorithm 2 and Algorithm 3 separately to both the row player and the column player leads to the output $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ such that $(\mathbb{E}[\bar{\mathbf{x}}], \mathbb{E}[\bar{\mathbf{y}}])$ is an ε-close NE. A similar δ-independent sample complexity guarantee can be obtained by adapting Theorem 6 and Theorem 8 to the dual LP.

F Proof of Theorem 4

Theorem 4 follows from the theory of simplex method for solving LP. We rewrite the LP V^{Prime} in Equation (10) in the standard form as

$$V^{\mathrm{Prime}} = \min_{\mathbf{x},\mathbf{s}\succeq 0, \mu\in\mathbb{R}} \mu \quad \text{s.t. } A^\top \mathbf{x} + \mathbf{s} - \mu \cdot \mathbf{e}^{m_2} = \mathbf{0},\ (\mathbf{e}^{m_1})^\top \mathbf{x} = 1, \tag{24}$$

where we introduce the slackness variables $\mathbf{s} \in \mathbb{R}^{m_2}$. The index set for the variables $(\mathbf{x}, \mathbf{s}, \mu)$ is given by $[m_1 + m_2 + 1]$ where the first m_1 coordinates correspond to $\mathbf{x}$, the next m_2 coordinates correspond to $\mathbf{s}$, and the last coordinate corresponds to μ. Denote by $\hat{\mathcal{I}}$ an optimal basis to the LP Equation (24). Then, we know that the corresponding optimal solution, denoted by $(\mathbf{x}^*, \mathbf{s}^*, \mu^*)$, can be derived as the *unique* solution to the linear system

$$A_{\hat{\mathcal{I}}_1,:}^\top \mathbf{x}_{\hat{\mathcal{I}}_1} + \mathbf{s} - \mu \cdot \mathbf{e}^{m_2} = \mathbf{0},\ (\mathbf{e}^{|\hat{\mathcal{I}}_1|})^\top \mathbf{x}_{\hat{\mathcal{I}}_1} = 1,\ \mathbf{x}_{\hat{\mathcal{I}}_1^c} = \mathbf{0},\ \mathbf{s}_{\hat{\mathcal{I}}_2^c} = \mathbf{0}, \tag{25}$$

where we denote by $\hat{\mathcal{I}}_1$ as $\hat{\mathcal{I}}$ restricted to the index sets of $\mathbf{x}$ and $\hat{\mathcal{I}}_2$ as $\hat{\mathcal{I}}$ restricted to the index sets of $\mathbf{s}$. Define $\hat{\mathcal{I}}_1^c = [m_1]\backslash\hat{\mathcal{I}}$ to be the complementary set of $\hat{\mathcal{I}}_1$ with respect to x and $\hat{\mathcal{I}}_2^c = [m_1+m_2]\backslash[m_1]\backslash\hat{\mathcal{I}}$ to be the complementary set of $\hat{\mathcal{I}}_2$ with respect to $\mathbf{s}$. Note that there are $m_2 + 1$ constraints in V^{Prime} in (24) . Then the basis set $\hat{\mathcal{I}}$ should contain $m_2 + 1$ elements, with one element being the one denoting the index for μ. This implies that $|\hat{\mathcal{I}}_1| + |\hat{\mathcal{I}}_2| = m_2$, which implies $|\hat{\mathcal{I}}_1| = |\hat{\mathcal{I}}_2^c|$. Then, we know that the linear system

$$A_{\hat{\mathcal{I}}_1,\hat{\mathcal{I}}_2}^\top \mathbf{x}_{\hat{\mathcal{I}}_1} + \mathbf{s}_{\hat{\mathcal{I}}_2} = \mu \cdot \mathbf{e}^{|\hat{\mathcal{I}}_2|},\ A_{\hat{\mathcal{I}}_1,\hat{\mathcal{I}}_2^c}^\top \mathbf{x}_{\hat{\mathcal{I}}_1} = \mu \cdot \mathbf{e}^{|\hat{\mathcal{I}}_2^c|},\ (\mathbf{e}^{|\hat{\mathcal{I}}_1|})^\top \mathbf{x}_{\hat{\mathcal{I}}_1} = 1, \tag{26}$$

also has a unique optimal solution, which is $(\mathbf{x}_{\hat{\mathcal{I}}_1}^*, \mathbf{s}_{\hat{\mathcal{I}}_2}, \mu^*)$. Otherwise, if the linear system Equation (26) has a different solution, then we append the solution with $\mathbf{x}_{\hat{\mathcal{I}}_1^c} = \mathbf{0}$ and $\mathbf{s}_{\hat{\mathcal{I}}_2^c} = \mathbf{0}$, and we obtain a different solution to Equation (25), which violates the uniqueness of the optimal solution $(\mathbf{x}^*, \mathbf{s}^*, \mu^*)$.

Moreover, the uniqueness of the optimal solution to Equation (26) implies that all the linear equations in Equation (26) are linearly independent of each other, since the number of equations and variables are equivalent to each other. Therefore, we know that the subset of the equations in Equation (26), which is shown in the following linear system

$$A^\top_{\hat{\mathcal{I}}_1, \hat{\mathcal{I}}_2^c} \mathbf{x}_{\hat{\mathcal{I}}_1} = \mu \cdot \mathbf{e}^{|\hat{\mathcal{I}}_2^c|}, \ (\mathbf{e}^{|\hat{\mathcal{I}}_1|})^\top \mathbf{x}_{\hat{\mathcal{I}}_1} = 1 \tag{27}$$

is also linearly independent of each other. Further note that the number of equations and variables in Equation (27) is equivalent to each other, we know that Equation (27) enjoys a unique optimal solution, which is given by $(\mathbf{x}^*_{\hat{\mathcal{I}}_1}, \mu^*)$. Therefore, we know that the square matrix

$$\hat{A} = \begin{bmatrix} A^\top_{\hat{\mathcal{I}}_1, \hat{\mathcal{I}}_2^c} & -\mathbf{e}^{|\hat{\mathcal{I}}_2^c|} \\ (\mathbf{e}^{|\hat{\mathcal{I}}_1|})^\top, & 0 \end{bmatrix} \tag{28}$$

is of full rank and all of its columns are linear independent. Further, we denote by $\mathcal{I}^*$ the support set of $\mathbf{x}^*$ such that $\mathbf{x}^*_{\mathcal{I}^*} \succ \mathbf{0}$ and $\mathbf{x}^*_{\mathcal{I}^{*c}} = \mathbf{0}$. Then, we know that $\mathcal{I}^* \subset \hat{\mathcal{I}}_1$ and the matrix

$$\hat{A}^* = \begin{bmatrix} A^\top_{\mathcal{I}^*, \hat{\mathcal{I}}_2^c} & -\mathbf{e}^{|\hat{\mathcal{I}}_2^c|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top, & 0 \end{bmatrix} \tag{29}$$

is of full column rank. We also know that $(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)$ can be expressed as a solution to the linear system

$$\hat{A}^* \cdot \begin{bmatrix} \mathbf{x}^*_{\mathcal{I}^*} \\ \mu^* \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \tag{30}$$

We now specify an index set $\mathcal{J}^* \subset \hat{\mathcal{I}}_2^c$ such that $\mathcal{I}^*$ and $\mathcal{J}^*$ satisfy the requirements. Note that the rank of a matrix is given by the largest rank of all its square submatrix. Therefore, we know that there exists a square submatrix of $\hat{A}^*$, denoted as A^*, such that the rank equals $|\mathcal{I}^*| + 1$. On the other hand, the matrix $\hat{A}^*$ has $|\mathcal{I}^*| + 1$ number of columns. Therefore, we know that the submatrix A^* is obtained from selecting $|\mathcal{I}^*| + 1$ number of linearly independent rows from the matrix $\hat{A}^*$ (note that for square matrix A^*, full row rank implies full rank of $|\mathcal{I}^*| + 1$). In the following claim, we show that there exists such a square submatrix A^* which contains the last row $[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0]$.

Claim 1 *There exists a square submatrix A^* of $\hat{A}^*$ such that A^* is of full rank (of rank $|\mathcal{I}^*| + 1$) and A^* contains the last row $[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0]$.*

Therefore, we know that the square matrix

$$A^* = \begin{bmatrix} A^\top_{\mathcal{I}^*, \mathcal{J}^*} & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top, & 0 \end{bmatrix} \tag{31}$$

is of full rank, and it holds that

$$A^* \cdot \begin{bmatrix} \mathbf{x}^*_{\mathcal{I}^*} \\ \mu^* \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \tag{32}$$

which also implies that $(\mathbf{x}^*_{\mathcal{I}^*, \mu^*})$ is the unique solution to the linear system Equation (32) with $\mathbf{x}^*_{\mathcal{I}^{*c}} = \mathbf{0}$. We also have $|\mathcal{I}^*| = |\mathcal{J}^*|$. Our proof is thus completed.

F.1 Proof of Claim 1

We show that any full rank square submatrix (with rank $|\mathcal{I}^*| + 1$) of $\hat{A}^*$ can be converted into a full rank square submatrix A^* (with rank $|\mathcal{I}^*| + 1$) containing the last row $[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0]$. Suppose that we have such a full rank square submatrix A' consisting of $|\mathcal{I}^*| + 1$ rows from $[A^\top_{\mathcal{I}^*, \hat{\mathcal{I}}_2^c}, -\mathbf{e}^{|\hat{\mathcal{I}}_2^c|}]$. We denote by the rows of A' as $\mathbf{a}_1, \dots, \mathbf{a}_{d+1}$, where we set $d = |\mathcal{I}^*|$. Then, since the matrix $\hat{A}^*$ is of rank $d+1$, we know that there exists a set of constants $\alpha_1, \dots, \alpha_{d+1}$ such that the last row can be expressed as

$$[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0] = \sum_{i=1}^{d+1} \alpha_i \cdot \mathbf{a}_i. \tag{33}$$

From comparing the last elements of the rows in Equation (33), we know that

$$0 = \sum_{i=1}^{d+1} \alpha_i. \tag{34}$$

Now suppose that $\alpha_{d+1} \neq 0$, we know that $\alpha_{d+1} = -\sum_{i=1}^{d} \alpha_i$. Therefore, we have

$$\mathbf{a}_{d+1} = \frac{1}{\alpha_{d+1}} \cdot \left([(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0] + \sum_{i=1}^{d} \alpha_i \cdot \mathbf{a}_i \right). \tag{35}$$

Since every row of $\hat{A}^*$ can be expressed as the linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_{d+1}$, it can also be expressed as the linear combination of $[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0]$ and $\mathbf{a}_1, \dots, \mathbf{a}_d$ by noting Equation (35). Therefore, we know that a square submatrix A^* consisting of $[(\mathbf{e}^{|\mathcal{I}^*|})^\top, 0]$ and $\mathbf{a}_1, \dots, \mathbf{a}_d$ is of the same rank as $\hat{A}^*$, which is $d+1$, and thus A^* is of full row rank, thus full rank.

G Proof of Theorem 5

We now condition on the event that

$$\mathcal{E} = \left\{ \left| \widehat{A}_{i,j} - A_{i,j} \right| \leq \mathrm{Rad}(N/m, \varepsilon/m), \ \ \forall i \in [m_1], j \in [m_2] \right\}, \tag{36}$$

where $\widehat{A}_{i,j}$ is constructed using N/m number of i.i.d samples with mean $A_{i,j}$, $m = m_1 m_2$, and $\mathrm{Rad}(N, \varepsilon) \triangleq \sqrt{\frac{\log(1/\varepsilon)}{2N}}$. According to Hoeffding's inequality, we know that this event $\mathcal{E}$ happens with probability at least $1 - \varepsilon$.

Given a fixed sample set $\mathcal{H}$ such that event $\mathcal{E}$ happens. We first bound the gap between $V_{\mathcal{I}}^{\mathrm{Prime}}$ and $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$, for any set $\mathcal{I}$. The result is formalized in the following claim.

Claim 2 *Conditioned on the event $\mathcal{E}$ Equation* (36) *happens, for any set $\mathcal{I}$, it holds that*

$$\left| V_{\mathcal{I}}^{\mathrm{Prime}} - \widehat{V}_{\mathcal{I}}^{\mathrm{Prime}} \right| \leq \mathrm{Rad}(N/m, \varepsilon/m). \tag{37}$$

We can also obtain the bound of the gap between the the objective values of the dual LPs, $V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ and $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$, for any sets $\mathcal{I}$ and $\mathcal{J}$. The bound is formalized in the following claim.

Claim 3 *Conditioned on the event $\mathcal{E}$ Equation* (36) *happens, for any set $\mathcal{I}$ and $\mathcal{J}$, it holds that*

$$\left| V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} - \widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} \right| \leq \mathrm{Rad}(N/m, \varepsilon/m). \tag{38}$$

Equipped with Claim 2 and Claim 3, we now proceed our proof. We first show that the output of Algorithm 2, denoted as $\mathcal{I}^*$ and $\mathcal{J}^*$, is indeed an optimal basis to the estimated LP $\widehat{V}^{\mathrm{Prime}}$ and its dual $\widehat{V}^{\mathrm{Dual}}$. We then show that as long as $N \geq N_0 = \frac{2m\log(2m/\epsilon)}{\delta^2}$, an optimal basis to $\widehat{V}^{\mathrm{Prime}}$ is an optimal basis to V^{Prime}.

Following the procedure in Algorithm 2, we identify an index set $\mathcal{I} \subset [m_1]$, where we set $x_i = 0$ for each $i \in \mathcal{I}^c$. Note that we cannot further delete one more i from the set $\mathcal{I}$ without changing the objective value according to our algorithm design. Thus, we know that our Algorithm 2 correctly identifies an index set $\mathcal{I}$, which is an optimal basis to the estimated LP $\widehat{V}^{\mathrm{Prime}}$.

Now given that $\mathcal{I}$ is an optimal basis to the LP $\widehat{V}^{\mathrm{Prime}}$, we know that

$$\widehat{V}^{\mathrm{Prime}} = \widehat{V}_{\mathcal{I}}^{\mathrm{Prime}} = \min_{\mathbf{x}_{\mathcal{I}} \succeq 0, \mu \in \mathbb{R}} \mu \quad \text{s.t. } \mu \cdot \mathbf{e}^{m_2} \succeq \widehat{A}_{\mathcal{I},:}^\top \mathbf{x}_{\mathcal{I}}, \ (\mathbf{e}^{|\mathcal{I}|})^\top \mathbf{x}_{\mathcal{I}} = 1. \tag{39}$$

Note that in the formulation of Equation (39), we simply discard the columns of the constraint matrix in the index set $\mathcal{I}$. Thus, if we denote by $\hat{\mathbf{x}}^*$ the optimal solution to $\widehat{V}^{\text{Prime}}$ corresponding to the basis $\mathcal{I}$, then one optimal solution to Equation (39) will just be $\hat{\mathbf{x}}^*_{\mathcal{I}}$. The dual of Equation (39) is

$$\widehat{V}^{\text{Dual}}_{\mathcal{I}} = \max_{\nu \geq 0, \mathbf{y} \succeq 0} \nu \quad \text{s.t. } \nu \cdot \mathbf{e}^{|\mathcal{I}|} \leq \widehat{A}_{\mathcal{I},:} \cdot \mathbf{y},\ (\mathbf{e}^{m_2})^\top \mathbf{y} = 1. \tag{40}$$

We now show that Algorithm 2 correctly identify an optimal basis of the estimated LP $\widehat{V}^{\text{Dual}}_{\mathcal{I}}$, denoted by $\mathcal{J}$. Note that following the procedure of Algorithm 2, we identify an index set $\mathcal{J} \subset [m_2]$, where we set $y_j = 0$ for each $j \in \mathcal{J}^c$. Similarly, we know that we cannot further delete one more j from the set $\mathcal{J}$ without changing the objective value according to our algorithm design. Thus, we know that our Algorithm 2 correctly identifies the index set $\mathcal{J}$, which is an optimal basis to the estimated LP $\widehat{V}^{\text{Dual}}_{\mathcal{I}}$, conditional on the event $\mathcal{E}$ Equation (36) happens.

It remains to show that the index sets $\mathcal{I}$ and $\mathcal{J}$ identified by our Algorithm 2 satisfy the conditions in Theorem 4, for the estimated LP $\widehat{V}^{\text{Prime}}$ and its dual $\widehat{V}^{\text{Dual}}$. Denote by $(\hat{\mathbf{x}}^*, \hat{\mathbf{y}}^*)$ an optimal primal-dual solution to $\widehat{V}^{\text{Prime}}$, corresponding to the optimal basis $\mathcal{I}$ and $\mathcal{J}$. Since we know that $\hat{\mathbf{y}}^*_{\mathcal{J}} > 0$, from the complementary slackness condition, the corresponding constraints in $\widehat{V}^{\text{Prime}}$ must be binding, i.e., $\widehat{A}^\top_{:,\mathcal{J}} \cdot \hat{\mathbf{x}}^* = \hat{\mu}^* \cdot \mathbf{e}$, where we denote by $(\hat{\mathbf{x}}^*, \hat{\mu}^*)$ the optimal solution to $\widehat{V}^{\text{Prime}}$ corresponding to the optimal basis $\mathcal{I}$ and $\mathcal{J}$. Moreover, note that $\hat{\mathbf{x}}^*_{\mathcal{I}^c} = 0$, we must have

$$\widehat{A}^\top_{:,\mathcal{J}} \cdot \hat{\mathbf{x}}^* = \widehat{A}^\top_{\mathcal{I},\mathcal{J}} \cdot \hat{\mathbf{x}}^*_{\mathcal{I}} = \hat{\mu}^* \cdot \mathbf{e}^{|\mathcal{J}|},$$

which proves the equations in condition Equation (12).

Next, we show that $|\mathcal{I}| = |\mathcal{J}|$, and $\widehat{A}_{\mathcal{I},\mathcal{J}}$ is a non-singular matrix. Note that the index set $\mathcal{J}$ is obtained from a basic optimal solution to the dual LP Equation (40), where there are $|\mathcal{I}|$ constraints. Thus, the matrix $\widehat{A}_{\mathcal{I},\mathcal{J}}$ can have at most $|\mathcal{I}|$ number of linearly independent columns, which implies that $|\mathcal{J}| \leq |\mathcal{I}|$, and the columns in the matrix $\widehat{A}_{\mathcal{I},\mathcal{J}}$ are linearly independent from each other. Now suppose $|\mathcal{I}| > |\mathcal{J}|$. Since the row and column ranks are equivalent, we know that at least one row of $\widehat{A}_{\mathcal{I},\mathcal{J}}$, denoted by i, can be expressed as the linear combination of the other rows. However, this means that we can further restrict $x_i = 0$ by removing i from the index set $\mathcal{I}$ without changing the objective value, which leads to a contradiction. Therefore, we know that $|\mathcal{I}| = |\mathcal{J}|$ and the square matrix $\widehat{A}_{\mathcal{I},\mathcal{J}}$ is of full rank.

Finally, it remains to show that $\mathcal{I}$ and $\mathcal{J}$ is also an optimal basis to the original LP V^{Prime} and its dual V^{Dual} as long as $N \geq N_0 = \frac{2m \log(2m/\epsilon)}{\delta^2}$. Suppose that $\mathcal{I}$ is not an optimal basis to the original LP V^{Prime}, then following the definition of $\delta = \min\{\delta_1, \delta_2\}$ with δ_1 and δ_2 given in Definition 5, we must have

$$V^{\text{Prime}}_{\mathcal{I}} > V^{\text{Prime}} + \delta.$$

On the other hand, from Equation (37) in Claim 2, we know that

$$|V^{\text{Prime}}_{\mathcal{I}} - \widehat{V}^{\text{Prime}}_{\mathcal{I}}| \leq \text{Rad}(N/m, \varepsilon/m),$$

and

$$|V^{\text{Prime}} - \widehat{V}^{\text{Prime}}| \leq \text{Rad}(N/m, \varepsilon/m).$$

However, we know that

$$\widehat{V}^{\text{Prime}}_{\mathcal{I}} = \widehat{V}^{\text{Prime}}$$

since $\mathcal{I}$ is an optimal basis to $\widehat{V}^{\text{Prime}}$. Therefore, we conclude that we have

$$V^{\text{Prime}} + 2\text{Rad}(N/m, \varepsilon/m) \geq V^{\text{Prime}}_{\mathcal{I}} > V^{\text{Prime}} + \delta. \tag{41}$$

However, the inequality Equation (41) contradicts with the condition that $\text{Rad}(N/m, \varepsilon/m) \leq \delta/2$ according to the definition of $\text{Rad}(N/m, \varepsilon/m)$. Therefore, we know that $\mathcal{I}$ is also an optimal basis to V^{Prime}.

We then show that $\mathcal{J}$ is an optimal basis to the dual LP $V^{\text{Dual}}_{\mathcal{I}}$. Suppose that $\mathcal{J}$ is not an optimal basis to the original dual LP $V^{\text{Dual}}_{\mathcal{I}}$, then following the definition of $\delta = \min\{\delta_1, \delta_2\}$ with δ_1 and δ_2 given in Definition 5, we must have

$$V^{\text{Dual}}_{\mathcal{I},\mathcal{J}} < V^{\text{Dual}}_{\mathcal{I}} - \delta.$$

On the other hand, from Equation (38) in Claim 3, we know that

$$|V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} - \widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}| \leq \mathrm{Rad}(N/m, \varepsilon/m),$$

and

$$|V_{\mathcal{I}}^{\mathrm{Dual}} - \widehat{V}_{\mathcal{I}}^{\mathrm{Dual}}| \leq \mathrm{Rad}(N/m, \varepsilon/m).$$

However, we know that

$$\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} = \widehat{V}_{\mathcal{I}}^{\mathrm{Dual}}$$

since $\mathcal{J}$ is an optimal basis to $\widehat{V}_{\mathcal{I}}^{\mathrm{Dual}}$. Therefore, we conclude that we have

$$V_{\mathcal{I}}^{\mathrm{Dual}} - 2\mathrm{Rad}(N/m, \varepsilon/m) \leq V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} \leq V_{\mathcal{I}}^{\mathrm{Dual}} - \delta. \tag{42}$$

However, the inequality Equation (42) contradicts with the condition that $\mathrm{Rad}(N/m, \varepsilon/m) \leq \delta/2$. Therefore, we know that $\mathcal{J}$ is also an optimal basis to $V_{\mathcal{I}}^{\mathrm{Dual}}$. We conclude that $\mathcal{I}$ and $\mathcal{J}$ is an optimal basis to the original LP V^{Prime} and its dual V^{Dual}, i.e., $\mathcal{I}$ and $\mathcal{J}$ satisfy the conditions described in Theorem 4 for the original LP V^{Prime} and V^{Dual}.

G.1 Proof of Claim 2

Denote by $\mathbf{x}^*$ an optimal solution to $V_{\mathcal{I}}^{\mathrm{Prime}}$. We now construct a feasible solution to $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$ based on $\mathbf{x}^*$. Note that conditional on the event $\mathcal{E}$ happens, we have that

$$\widehat{A}^\top \mathbf{x}^* \preceq A^\top \mathbf{x}^* + \|\mathbf{x}^*\|_1 \cdot \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_2} \preceq \mu^* \cdot \mathbf{e}^{m_2} + \cdot\mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_2}. \tag{43}$$

Therefore, we know that $(\mathbf{x}^*, \mu^* + \mathrm{Rad}(N/m, \varepsilon/m))$ is a feasible solution to the LP $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$ defined in Equation (13), meaning that

$$\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}} \leq \hat{\mu} = \mu^* + \mathrm{Rad}(N/m, \varepsilon/m) = V_{\mathcal{I}}^{\mathrm{Prime}} + \mathrm{Rad}(N/m, \varepsilon/m). \tag{44}$$

On the other hand, denote by $\hat{\mathbf{x}}^*$ and $\hat{\mu}^*$ certain optimal solution to $\widehat{V}_{\mathcal{I}}^{\mathrm{Prime}}$. We know that

$$A^\top \hat{\mathbf{x}}^* \preceq \widehat{A}^\top \hat{\mathbf{x}}^* + \|\hat{\mathbf{x}}^*\|_1 \cdot \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_2} \preceq \hat{\mu}^* \cdot \mathbf{e}^{m_2} + \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_2}. \tag{45}$$

Therefore, we know that $(\hat{\mathbf{x}}^*, \hat{\mu}^* + \mathrm{Rad}(N/m, \varepsilon/m))$ is a feasible solution to the LP $V_{\mathcal{I}}^{\mathrm{Prime}}$, meaning that

$$V_{\mathcal{I}}^{\mathrm{Prime}} \leq \hat{\mu}^* + \mathrm{Rad}(N/m, \varepsilon/m) = \widehat{V}_{\mathcal{I}}^{\mathrm{Prime}} + \mathrm{Rad}(N/m, \varepsilon/m). \tag{46}$$

Our proof is completed from Equation (44) and Equation (46).

G.2 Proof of Claim 3

Denote by $(\mathbf{y}^*, \nu^*)$ an optimal solution to $V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$. Similar to the proof for Claim 3, we now construct a feasible solution to $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ based on $\mathbf{y}^*$. Note that conditional on the event $\mathcal{E}$ happens, we have that

$$\widehat{A}_{\mathcal{I},:} \cdot \mathbf{y}^* \succeq A_{\mathcal{I},:} \cdot \mathbf{y}^* - \|\mathbf{y}^*\|_1 \cdot \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_1} \succeq \nu^* - \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_1}. \tag{47}$$

Therefore, $(\mathbf{y}^*, \nu^* - \mathrm{Rad}(N/m, \varepsilon/m))$ forms a feasible solution to $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$. Thus, we know that

$$\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} \geq \nu^* - \mathrm{Rad}(N/m, \varepsilon/m) = V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} - \mathrm{Rad}(N/m, \varepsilon/m). \tag{48}$$

On the other hand, denote by $\hat{\mathbf{y}}^*$ and $\hat{\nu}^*$ an optimal solution to $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$. We know that

$$A_{\mathcal{I},:} \cdot \hat{\mathbf{y}}^* \succeq \widehat{A}_{\mathcal{I},:} \cdot \hat{\mathbf{y}}^* - \|\hat{\mathbf{y}}^*\|_1 \cdot \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_1} \succeq \hat{\nu}^* - \mathrm{Rad}(N/m, \varepsilon/m) \cdot \mathbf{e}^{m_1}. \tag{49}$$

Therefore, $(\hat{\mathbf{y}}^*, \hat{\nu}^* - \mathrm{Rad}(N/m, \varepsilon/m))$ forms a feasible solution to $V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ and

$$V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} \geq \hat{\nu}^* - \mathrm{Rad}(N/m, \varepsilon/m) = \widehat{V}_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}} - \mathrm{Rad}(N/m, \varepsilon/m). \tag{50}$$

Our proof is thus completed from Equation (48) and Equation (50).

H Proof of Theorem 6

Consider a fixed N and we condition on the following high-probability event

$$\mathcal{E} = \left\{ \left| \widehat{A}_{i,j} - A_{i,j} \right| \leq \mathrm{Rad}(N/m, \varepsilon/m), \ \ \forall i \in [m_1], j \in [m_2] \right\}, \tag{51}$$

which happens with probability at least $1 - \varepsilon$. We denote by $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$ the output of Algorithm 2. From the implementation of Algorithm 2, we know that $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$ is an optimal basis to the LP $\widehat{V}^{\mathrm{Prime}}$ and $\widehat{V}^{\mathrm{Dual}}_{\hat{\mathcal{I}}}$. Therefore, we know that

$$\widehat{V}^{\mathrm{Prime}} = \widehat{V}^{\mathrm{Prime}}_{\mathcal{I}} = \widehat{V}^{\mathrm{Dual}}_{\hat{\mathcal{I}},\hat{\mathcal{J}}}.$$

On the other hand, according to Claim 2 and Claim 3, we know that

$$|V^{\mathrm{Prime}} - \widehat{V}^{\mathrm{Prime}}| \leq \mathrm{Rad}(N/m, \varepsilon/m) \ \text{ and } \ |V^{\mathrm{Dual}}_{\hat{\mathcal{I}},\hat{\mathcal{J}}} - \widehat{V}^{\mathrm{Dual}}_{\hat{\mathcal{I}},\hat{\mathcal{J}}}| \leq \mathrm{Rad}(N/m, \varepsilon/m).$$

Therefore, we have

$$|V^{\mathrm{Prime}} - V^{\mathrm{Prime}}_{\hat{\mathcal{I}}}| \leq 2 \cdot \mathrm{Rad}(N/m, \varepsilon/m) \ \text{ and } \ |V^{\mathrm{Prime}} - V^{\mathrm{Dual}}_{\hat{\mathcal{I}},\hat{\mathcal{J}}}| \leq 2 \cdot \mathrm{Rad}(N/m, \varepsilon/m),$$

which completes our proof of Equation (16).

We now denote by $\hat{\mathbf{x}}$ the optimal primal-dual solution to the LP $\widehat{V}^{\mathrm{Prime}}$, corresponding to the optimal basis $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$. Therefore, it holds that

$$\hat{\mathbf{x}}_{\hat{\mathcal{I}}^c} = 0, \tag{52}$$

and $\hat{\mathbf{x}}_{\hat{\mathcal{I}}}$ is the solution to

$$\begin{bmatrix} \widehat{A}^{\top}_{\hat{\mathcal{I}},\hat{\mathcal{J}}} & -\mathbf{e}^{|\hat{\mathcal{J}}|} \\ (\mathbf{e}^{|\hat{\mathcal{I}}|})^{\top} & 0 \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{x}}_{\hat{\mathcal{I}}} \\ \hat{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \tag{53}$$

We now apply the perturbation analysis of linear equation to bound the distance between $\hat{\mathbf{x}}_{\hat{\mathcal{I}}}$ and $\mathbf{x}^*_{\hat{\mathcal{I}}}(\hat{\mathcal{I}}, \hat{\mathcal{J}})$ defined in Equation (15). We denote by $\hat{\sigma}_{\min}$ the smallest absolute value of the eigenvalues and $\hat{\sigma}_{\max}$ the largest absolute value of the eigenvalues of the matrix

$$\begin{bmatrix} \widehat{A}^{\top}_{\hat{\mathcal{I}},\hat{\mathcal{J}}} & -\mathbf{e}^{|\hat{\mathcal{J}}|} \\ (\mathbf{e}^{|\hat{\mathcal{I}}|})^{\top} & 0 \end{bmatrix}.$$

Since $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$ is the output of Algorithm 2, we know that $\hat{\sigma}_{\min} > 0$. The perturbation of the matrix is denoted as

$$\Delta A = A_{\hat{\mathcal{I}},\hat{\mathcal{J}}} - \widehat{A}_{\hat{\mathcal{I}},\hat{\mathcal{J}}}.$$

Then, it holds that

$$\|\Delta A\|_2 \leq d \cdot \mathrm{Rad}(N/m, \varepsilon/m). \tag{54}$$

We now classify into two situations.

If $\mathrm{Rad}(N/m, \varepsilon/m) > \frac{\hat{\sigma}_{\min}}{2d}$, we know that $d \cdot \widehat{\kappa} \cdot \mathrm{Rad}(N/m, \varepsilon/m) \geq \frac{1}{2}$, meaning that the suboptimality gap is $\mathcal{O}(d \cdot \widehat{\kappa}\mathrm{Rad}(N/m, \varepsilon/m))$ since it is at most $\mathcal{O}(1)$.

Otherwise, if $\mathrm{Rad}(N/m, \varepsilon/m) \leq \frac{\hat{\sigma}_{\min}}{2d}$, we know that

$$\|\Delta A\|_2 \leq d \cdot \mathrm{Rad}(N/m, \varepsilon/m) \leq \frac{\hat{\sigma}_{\min}}{2}. \tag{55}$$

Following standard perturbation analysis of linear equations (Theorem 1 of [14]), we can obtain that

$$\frac{\|(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu}) - (\mathbf{x}^*_{\hat{\mathcal{I}}}(\hat{\mathcal{I}}, \hat{\mathcal{J}}), \mu^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}))\|_2}{\|(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu})\|_2} \leq \frac{\hat{\sigma}_{\max}/\hat{\sigma}_{\min}}{1 - \frac{\|\Delta A\|_2}{\hat{\sigma}_{\min}}} \cdot \left(\frac{\|\Delta A\|_2}{\hat{\sigma}_{\max}} \right) \leq 2 \cdot \|\Delta A\|_2 / \hat{\sigma}_{\min}. \tag{56}$$

Since $(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu})$ is the solution to the linear equations in Equation (53), we know that

$$\|\widehat{A}^{\top}_{\hat{\mathcal{I}},\hat{\mathcal{J}}} \hat{\mathbf{x}}_{\hat{\mathcal{I}}}\|_\infty \leq d$$

and thus $|\hat{\mu}| \leq d$. Therefore, we know that

$$\|(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu})\|_2 \leq 2d. \tag{57}$$

Combining Equation (57) and Equation (56), we can obtain that

$$\|(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu}) - (\mathbf{x}^*_{\hat{\mathcal{I}}}(\hat{\mathcal{I}}, \hat{\mathcal{J}}), \mu^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}))\|_2 \leq \frac{4d^2}{\hat{\sigma}_{\min}} \cdot \mathrm{Rad}(N/m, \varepsilon/m). \tag{58}$$

From the feasibility of $(\hat{\mathbf{x}}_{\hat{\mathcal{I}}}, \hat{\mu})$ to the LP $\widehat{V}^{\mathrm{Prime}}$, we know that

$$\hat{\mu} \cdot \mathbf{e}^{m_2} \succeq \widehat{A}^\top \cdot \hat{\mathbf{x}}, \tag{59}$$

which implies that

$$\begin{aligned} A^\top \mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) &= \widehat{A}^\top \cdot \hat{\mathbf{x}} + \widehat{A}^\top \cdot (\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) - \hat{\mathbf{x}}) + \Delta A^\top \cdot \mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) \\ &\preceq \left(\hat{\mu} + \|\widehat{A}^\top(\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) - \hat{\mathbf{x}})\|_\infty + \|\Delta A^\top \mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})\|_\infty\right) \mathbf{e}^{m_2} \\ &\preceq \left(\hat{\mu} + \|\widehat{A}^\top(\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) - \hat{\mathbf{x}})\|_2 + \|\Delta A^\top \mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})\|_\infty\right) \mathbf{e}^{m_2} \\ &\preceq \left(\hat{\mu} + \frac{4d^2 \widehat{\sigma}_{\max}}{\widehat{\sigma}_{\min}} \cdot \mathrm{Rad}(N/m, \varepsilon/m) + \mathrm{Rad}(N/m, \varepsilon/m)\right) \mathbf{e}^{m_2}. \end{aligned} \tag{60}$$

We conclude that we can set

$$\mu \triangleq \hat{\mu} + \mathrm{Rad}(N/m, \varepsilon/m) + \frac{4d^2 \hat{\sigma}_{\max}}{\hat{\sigma}_{\min}} \cdot \mathrm{Rad}(N/m, \varepsilon/m)$$

such that $(\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}), \mu)$ is a feasible solution to V^{Prime}. Therefore, we know that the sub-optimality gap for the solution $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$ is upper bounded by $\mathcal{O}\left(d^2 \cdot \widehat{\kappa} \cdot \mathrm{Rad}(N/m, \varepsilon/m)\right)$.

I Proof of Theorem 7

For a fixed dataset $\mathcal{H}^n = \{(i_k, j_k, \tilde{A}_{k,i_k,j_k})\}_{k=1}^n$, we denote by $\widehat{A}(\mathcal{H}^n)$ the sample-average estimation of A constructed from the dataset $\mathcal{H}^n$. Let n denote the n-th round of iteration in Algorithm 3. In the following, we condition on the event that Algorithm 2 has successfully identified the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$, which happens with probability at least $1 - \varepsilon$ from Theorem 5. We consider the stochastic process $\tilde{\mathbf{a}}^n$ defined in Equation (20). For a fixed $\eta > 0$ which will be specified later, we define sets

$$\mathcal{A} = \{\tilde{\mathbf{a}}' \in \mathbb{R}^{|\mathcal{J}^*|} : \tilde{a}'_j \in [-\eta, +\eta], \forall j \in \mathcal{J}^*\}, \tag{61}$$

Since $\tilde{\mathbf{a}}^1 = \mathbf{0}$, we have $\tilde{\mathbf{a}}^1 \in \mathcal{A}$. Then, we show that $\tilde{\mathbf{a}}^n$ "behave well" as long as they stay in the region $\mathcal{A}$ for a sufficiently long time, where $\tilde{\mathbf{a}}^n = \frac{\mathbf{a}^n}{N-n+1}$. To this end, we define a stopping time

$$\tau = \min_{n \in [N]} \{\tilde{\mathbf{a}}^n \notin \mathcal{A}\}. \tag{62}$$

The first lemma shows that when n is large enough but smaller than the stopping time τ, we have $\|(\mathbf{x}^n, \mu^n)\|_1 \leq L$ with $L = 4$. The proof is deferred to Section I.1.

Lemma 2 *There exist a constants N_0' such that when $N_0' \leq n \leq \tau$, it holds that $\|(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)\|_1 \leq L$ with probability at least $1 - 2\varepsilon$, where $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ denotes the solution to Equation* (18)*, τ is defined in Equation* (62) *with*

$$\eta = \frac{\sigma}{8d^{3/2} \cdot \|A_{\mathcal{I}^*, \mathcal{J}^*}\|_2}. \tag{63}$$

Specifically, N_0' is given as follows

$$N_0' = \frac{128d^5}{\sigma^2} \cdot \log(2d^2/\varepsilon). \tag{64}$$

The next lemma bounds $\mathbb{E}[N - \tau]$ with the proof deferred to Section I.2.

Lemma 3 *Let the stopping time τ be defined in Equation* (62) *with $\eta = \frac{\sigma}{8d^{3/2}\cdot\|A_{\mathcal{I}^*,\mathcal{J}^*}\|_2}$. It holds that*

$$\mathbb{E}[N-\tau] \le N_0' + 1 + 2\cdot\exp(-\eta^2/(8d^4)),$$

where N_0' is given in Equation (64)*, as long as*

$$N \ge \frac{1024d^3\cdot\|A_{\mathcal{I}^*,\mathcal{J}^*}\|_2^2\cdot\log(4d^2/\varepsilon)}{\sigma^2}. \tag{65}$$

In addition, for any N' such that $N_0' \le N' \le N$, it holds that

$$\Pr[\tau \le N'] \le \frac{\eta^2}{4d^4}\cdot\exp\left(-\frac{\eta^2(N-N')}{8d^4}\right) + \varepsilon. \tag{66}$$

Now we are ready to prove Theorem 7. From the definition of the stopping time τ in Equation (62), we know that for each $j \in \mathcal{J}^*$, it holds

$$a_j^{\tau-1} \in [-(N-\tau+1)\cdot\eta, (N-\tau+1)\cdot\eta].$$

Since (i_n, j_n) is uniformly drawn from $\mathcal{I}^*\times\mathcal{J}^*$ and is independent of the randomness of $\mathbf{x}^n$, we know that

$$\mathbb{E}_{\tilde{A}_n,i_n,j_n}\left[|\mathcal{J}^*|\cdot|\mathcal{I}^*|\cdot\tilde{A}_{n,i_n,j_n}\cdot x_{i_n}^n\cdot\mathbf{h}_{j_n}\right] = \sum_{i=1}^{|\mathcal{I}^*|}\sum_{j=1}^{|\mathcal{J}^*|} A_{ij}x_i^n\cdot\mathbf{h}_j = A_{\mathcal{I}^*,\mathcal{J}^*}^\top\mathbf{x}^n.$$

Then, following the update in Equation (19), we know that

$$\mathbb{E}[\mathbf{a}^{N+1}] = \mathbb{E}\left[\mathbf{a}^{\tau-1}\right] - \sum_{n=\tau}^{N} A_{\mathcal{I}^*,\mathcal{J}^*}^\top\cdot\mathbb{E}[\mathbf{x}_{\mathcal{I}^*}^n] + \sum_{n=\tau}^{N}\mathbb{E}[\mu^n]\cdot\mathbf{e}^{|\mathcal{J}^*|},$$

Thus, we have that

$$\left|\mathbb{E}[\mathbf{a}^{N+1}]\right| \le \left|\mathbb{E}[\mathbf{a}^{\tau-1}]\right| + \left|\sum_{n=\tau}^{N}\left(A_{\mathcal{I}^*,\mathcal{J}^*}^\top\cdot\mathbb{E}[\mathbf{x}_{\mathcal{I}^*}^n] - \mathbb{E}[\mu^n]\cdot\mathbf{e}^{|\mathcal{J}^*|}\right)\right|.$$

Following the definition of the stop time τ in (62), since it is easy to show that $\eta \le 1$, we know that

$$\left\|\mathbb{E}[\mathbf{a}^{\tau-1}]\right\|_\infty \le \mathbb{E}[N-\tau+2].$$

Also, from the boundedness that $\|(\mathbf{x}^n,\mu^n)\|_1 \le L$, we know that

$$\left\|\sum_{n=\tau}^{N}\left(A_{\mathcal{I}^*,\mathcal{J}^*}^\top\cdot\mathbb{E}[\mathbf{x}_{\mathcal{I}^*}^n] - \mathbb{E}[\mu^n]\cdot\mathbf{e}^{|\mathcal{J}^*|}\right)\right\|_\infty \le L\cdot(N-\tau).$$

As a result, for each $j \in \mathcal{J}^*$, it holds that

$$\left|\mathbb{E}[a_j^{N+1}]\right| \le (L+1)\cdot\mathbb{E}[N-\tau] + 2 \le (L+1)\cdot N_0' + 2(L+1)\cdot\exp(-\eta^2/(8d^4)) + 2, \tag{67}$$

where the second inequality holds because of the upper bound on $\mathbb{E}[N-\tau]$ established in Lemma 3. We further note that following the update in Equation (19), we have that

$$\begin{aligned}\mathbb{E}[\mathbf{a}^{N+1}] &= -\sum_{n=1}^{N} A_{\mathcal{I}^*,\mathcal{J}^*}^\top\cdot\mathbb{E}[\mathbf{x}_{\mathcal{I}^*}^n] + \sum_{n=1}^{N}\mathbb{E}[\mu^n]\cdot\mathbf{e}^{|\mathcal{J}^*|} \\ &= -N\cdot A_{\mathcal{I}^*,\mathcal{J}^*}^\top\cdot\mathbb{E}[\bar{\mathbf{x}}_{\mathcal{I}^*}] + N\cdot\mathbb{E}[\bar{\mu}]\cdot\mathbf{e}^{|\mathcal{J}^*|},\end{aligned}$$

where $\bar{\mu} = \frac{1}{N}\cdot\sum_{n=1}^N \mu^n$ and $\bar{\mathbf{x}} = \frac{1}{N}\cdot\sum_{n=1}^N \mathbf{x}^n$. Thus, we conclude that it holds

$$\begin{bmatrix} A_{\mathcal{I}^*,\mathcal{J}^*}^\top & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top & 0 \end{bmatrix}\cdot\begin{bmatrix}\mathbb{E}[\bar{\mathbf{x}}_{\mathcal{I}^*}] \\ \mathbb{E}[\bar{\mu}]\end{bmatrix} = \begin{bmatrix}\frac{\mathbb{E}[\mathbf{a}^{N+1}]}{N} \\ 1\end{bmatrix}, \tag{68}$$

We now compare the linear equations Equation (68) to Equation (17) to bound the gap between $\bar{\mathbf{x}}$ and $\mathbf{x}^*$, where $\mathbf{x}_{\mathcal{I}^c}^* = \bar{\mathbf{x}}_{\mathcal{I}^c} = \mathbf{0}$. Following standard perturbation analysis of linear equations (Theorem 1 of [14]), we know

$$\|\mathbf{x}^* - \mathbb{E}[\bar{\mathbf{x}}]\|_2 \le \kappa(A^*)\cdot\frac{\left\|\mathbb{E}[\mathbf{a}^{N+1}]\right\|_2}{N}, \tag{69}$$

where A^* is defined in Equation (23). We now plug in the formulation of N_0' that satisfies the conditions in Equation (64). We have that

$$\|\mathbf{x}^* - \mathbb{E}[\bar{\mathbf{x}}]\|_2 \leq O\left(\sqrt{d}\kappa(A^*) \cdot \left(\frac{d^5}{\sigma^2} + 1\right) \cdot \log(2d^2/\varepsilon) \cdot \frac{1}{N}\right). \tag{70}$$

In order to translate the bound in Equation (70) into the sample complexity bound, we let the right hand sides of Equation (70) equal ε, and we have that

$$N = O\left(d^{1/2} \cdot \kappa(A^*) \cdot \left(1 + \frac{d^5}{\sigma^2}\right) \cdot \frac{\log(1/\varepsilon)}{\varepsilon}\right).$$

Further upper bounding $\kappa(A^*)$ by $\frac{d}{\sigma}$ and combining with the complexity bound presented in Theorem 5 completes our proof.

I.1 Proof of Lemma 2

Denote by $(\mathbf{x}^*, \mu^*)$ and $(\mathbf{y}^*, \nu^*)$ the optimal primal-dual solution corresponding to the optimal basis $\mathcal{I}^*$ and $\mathcal{J}^*$. Then, it holds that

$$A^\top_{\mathcal{I}^*, \mathcal{J}^*} \cdot \mathbf{x}^*_{\mathcal{I}^*} = \mu^* \cdot \mathbf{e}^{|\mathcal{J}^*|}. \tag{71}$$

We compare $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ with $(\mathbf{x}^*, \mu^*)$ and when n is large enough. In the remaining proof, we focus on comparing $(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)$ with $(\mathbf{x}^*, \mu^*)$.

Note that when $n \geq N_0$, $\tilde{\mathbf{x}}^n$ is the solution to the following linear equations

$$\begin{bmatrix} \widehat{A}^\top_{\mathcal{I}^*, \mathcal{J}^*}(\mathcal{H}^n), & -\mathbf{e}^{|\mathcal{J}^*|} \\ (\mathbf{e}^{|\mathcal{I}^*|})^\top, & 0 \end{bmatrix} \cdot \begin{bmatrix} \tilde{\mathbf{x}}^n_{\mathcal{I}^*} \\ \tilde{\mu}^n \end{bmatrix} = \begin{bmatrix} \mathbf{a}^n/(N-n+1) \\ 1 \end{bmatrix}. \tag{72}$$

When $n \leq \tau$, we know that

$$\left|\frac{\mathbf{a}^n}{N-n+1}\right| \leq \eta. \tag{73}$$

Now, we denote by $\mathbb{I}\{i_q = i, j_q = j\}$ an indicator function of whether $i_q = i$ and $j_q = j$, and we denote by

$$n_{i,j} \triangleq \sum_{q=1}^{n} \mathbb{I}\{i_q = i, j_q = j\},$$

the number of samples of the matrix entry $A_{i,j}$. From the uniform sampling rule and standard Hoeffding bound, noting $\mathbb{E}[n_{i,j}] = n/d^2$, we know that for all

$$P[n_{i,j} \geq n/(2d^2)] = 1 - P(n_{i,j} < n/(2d^2)) \geq 1 - \exp(-n/(2d^4)).$$

We now condition on the event $\mathcal{E}'$ happens, which is defined as

$$\mathcal{E}' = \left\{n_{i,j} \geq n/(2d^2), \forall i \in \mathcal{I}^*, \forall j \in \mathcal{J}^*\right\}.$$

From the union bound, we know that

$$P[\mathcal{E}'] \geq 1 - d^2 \cdot \exp(-n/(2d^4)).$$

Conditional on the event $\mathcal{E}'$ happens, we know that the absolute value of each element of $\widehat{A}_{\mathcal{I}^*, \mathcal{J}^*}(\mathcal{H}^n) - A_{\mathcal{I}^*, \mathcal{J}^*}$ is upper bounded by $\mathrm{Rad}(n/(2d^2), \varepsilon/d^2)$, given that the following event

$$\mathcal{E} = \left\{\left|\frac{1}{n_{i,j}} \cdot \sum_{q=1}^{n_{i,j}} A_q(i,j) \cdot \mathbb{I}\{i_q = i, j_q = j\} - A_{i,j}\right| \leq \mathrm{Rad}(n/(2d^2), \varepsilon/d^2),\ \ \forall i \in \mathcal{I}^*, j \in \mathcal{J}^*\right\} \tag{74}$$

is assumed to happen (it holds with probability at least $1-\varepsilon$ following standard Hoeffding bound). We now bound the distance between the solutions to Equation (71) and Equation (72). The perturbation of the matrix is denoted as

$$\Delta A = A_{\mathcal{I}^*, \mathcal{J}^*} - \widehat{A}_{\mathcal{I}^*, \mathcal{J}^*}(\mathcal{H}^n).$$

It holds that

$$\|\Delta A\|_2 \leq d \cdot \mathrm{Rad}(n/(2d^2), \varepsilon/d^2). \tag{75}$$

Therefore, as long as

$$\|\Delta A\|_2 \le d \cdot \mathrm{Rad}(n/(2d^2), \varepsilon/d^2) \le \frac{1}{2 \cdot \|(A_{\mathcal{I}^*,\mathcal{J}^*})^{-1}\|_2} = \frac{\sigma}{2}, \tag{76}$$

following standard perturbation analysis of linear equations (Theorem 1 of [14]), we have that

$$\begin{aligned}\frac{\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_2}{\|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_2} &\le \frac{\kappa(A^*)}{1 - \kappa(A^*) \cdot \frac{\|\Delta A\|_2}{\|A^*\|_2}} \cdot \left(\frac{\|\Delta A\|_2}{\|A^*\|_2} + d \cdot \eta \right) \\ &\le 2 \cdot \kappa(A^*) \cdot \left(\frac{\|\Delta A\|_2}{\|A^*\|_2} + d \cdot \eta \right),\end{aligned} \tag{77}$$

where $\kappa(A^*)$ denotes the conditional number of A^*. We now want to propose conditions on $\|\Delta A\|_2$ and η such that the right hand side of (77) is small enough such that we can show the l_1 norm of $(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)$ is at most $L/2$. Together with the upper bound that $\|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le L/2$, our proof will be completed.

We note that $\|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_2 \le \|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le \frac{L}{2}$. Then, (77) can be rewritten as

$$\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_2 \le L \cdot \kappa(A^*) \cdot \left(\frac{\|\Delta A\|_2}{\|A^*\|_2} + d \cdot \eta \right),$$

which leads to

$$\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le \sqrt{d+1} \cdot \|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_2 \le 2\sqrt{d} \cdot L \cdot \kappa(A^*) \cdot \left(\frac{\|\Delta A\|_2}{\|A^*\|_2} + d \cdot \eta \right),$$

where last inequality follows from $\sqrt{d+1} \le 2\sqrt{d}$. We now set the condition on n such that n satisfies the condition in Equation (76) and the following condition

$$2\sqrt{d} \cdot L \cdot \kappa(A^*) \cdot \frac{\|\Delta A\|_2}{\|A^*\|_2} \le 2L \cdot d\sqrt{d} \cdot \frac{\mathrm{Rad}(n/(2d^2), \varepsilon/d^2)}{\sigma} \le \frac{L}{4}, \tag{78}$$

we also set the condition on η such that

$$2L \cdot \kappa(A^*) \cdot d\sqrt{d} \cdot \eta \le \frac{L}{4}. \tag{79}$$

Clearly, as long as conditions (76), (78), and (79) are satisfied, we have that

$$\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le 2\sqrt{d} \cdot L \cdot \kappa(A^*) \cdot \left(\frac{\|\Delta A\|_2}{\|A^*\|_2} + d \cdot \eta \right) \le L/2,$$

which will complete our proof.

We now set the condition $n \ge N_0'$ with N_0' given by

$$n \ge N_0' := \frac{128 d^5}{\sigma^2} \cdot \log(2d^2/\varepsilon). \tag{80}$$

Then, the conditions Equation (76) and Equation (78) are satisfied. Moreover, we set

$$\eta := \frac{\sigma}{8 d^{3/2} \cdot \|A^*\|_2}. \tag{81}$$

Then, the condition (79) is satisfied. Therefore, with the conditions (80) and (81), we know that

$$\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le L/2.$$

Moreover, note that $\|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le L/2$, we have that

$$\|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n)\|_1 \le \|(\tilde{\mathbf{x}}^n_{\mathcal{I}^*}, \tilde{\mu}^n) - (\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 + \|(\mathbf{x}^*_{\mathcal{I}^*}, \mu^*)\|_1 \le L.$$

Our proof is thus completed.

I.2 Proof of Lemma 3

Now we fix an arbitrary $j \in \mathcal{J}^*$. For any $N_0' \le N' \le N$, it holds that

$$\tilde{a}_j^{N'} - \tilde{a}_j^{N_0'} = \sum_{n=N_0'}^{N'-1} (\tilde{a}_j^{n+1} - \tilde{a}_j^n).$$

We define $\xi_j^n = \tilde{a}_j^{n+1} - \tilde{a}_j^n$. Then, we have

$$\tilde{a}_j^{N'} - \tilde{a}_j^{N_0'} = \sum_{n=N_0'}^{N'-1} (\xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{H}^n]) + \sum_{n=N_0'}^{N'-1} \mathbb{E}[\xi_j^n | \mathcal{H}^n],$$

where $\mathcal{H}^n$ denotes the filtration of information up to step n. Note that due to the update in Equation (21), we have

$$\xi_j^n = \frac{\tilde{a}_j^n - |\mathcal{J}^*| \cdot |\mathcal{I}^*| \cdot \tilde{A}_{n,i_n,j_n} \cdot x_{i_n}^n \cdot 1_{\{j=j_n\}} + \mu^n}{N-n},$$

where $1_{\{j=j_n\}}$ is an indicator function of whether $j = j_n$. Then, it holds that

$$|\xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{H}^n]| \le \frac{d^2}{N-n} \tag{82}$$

where the inequality follows from the fact that the value of $\tilde{a}_j^n$ is deterministic given the filtration $\mathcal{H}^n$ and we have $\|(\mathbf{x}^n, \mu^n)\|_1 \le L$ for any n, as well as $|\mathcal{J}^*| = |\mathcal{I}^*| = d$. Note that

$$\{\xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{H}^n]\}_{\forall n=\bar{N}_0,\dots,N'}$$

forms a martingale difference sequence. Following Hoeffding's inequality, for any $N'' \le N'$ and any $b > 0$, it holds that

$$\begin{aligned} P\left[ ] \left| \sum_{n=N_0'}^{N''} (\xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{H}^n]) \right| \ge b \right] &\le 2\exp\left(-\frac{b^2}{2 \cdot \sum_{n=N_0'}^{N''} d^4/(N-n)^2} \right) \\ &\le 2\exp\left(-\frac{b^2 \cdot (N - N'')}{d^4} \right). \end{aligned}$$

Therefore, we have that

$$\begin{aligned} &P\left[\left| \sum_{n=N_0'}^{N''} (\xi_j^n - \mathbb{E}[\xi_j^n | \mathcal{H}^n]) \right| \ge b \text{ for some } N_0' \le N'' \le N' \right] \\ &\le \sum_{N''=N_0'}^{N'} 2\exp\left(-\frac{b^2 \cdot (N-N'')}{2d^4} \right) \le \frac{b^2}{d^4} \cdot \exp\left(-\frac{b^2 \cdot (N-N')}{2d^4} \right) \end{aligned} \tag{83}$$

holds for any $b > 0$.

We now bound the probability that $\tau > N'$ for one particular N' such that $N_0' \le N' \le N$. Suppose that $N' \le \tau$, then, from Lemma 2, for each $n \le N'$, we know that $\|(\tilde{\mathbf{x}}^n, \tilde{\mu}^n)\|_1 \le L$ and therefore $\mathbf{x}^n = \tilde{\mathbf{x}}^n$, $\mu^n = \tilde{\mu}^n$ as the solution to Equation (18). We have

$$\tilde{a}_j^n = \widehat{A}_{\mathcal{I}^*,j}^\top(\mathcal{H}^n) \cdot \mathbf{x}_{\mathcal{I}^*}^n - \mu^n.$$

Now, we denote by $\mathbb{I}\{i_q = i, j_q = j\}$ an indicator function of whether $i_q = i$ and $j_q = j$, and we denote by

$$n_{i,j} \triangleq \sum_{q=1}^{n} \mathbb{I}\{i_q = i, j_q = j\},$$

the number of noisy samples for $A_{i,j}$ till iteration n. From the uniform sampling rule and standard Hoeffding bound, noting $\mathbb{E}[n_{i,j}] = n/d^2$, we know that

$$P[n_{i,j} \ge n/(2d^2)] = 1 - P[n_{i,j} < n/(2d^2)] \ge 1 - \exp(-n/(2d^4)).$$

We now condition on the event $\mathcal{E}'_n$ happens, which is defined as

$$\mathcal{E}'_n = \left\{n_{i,j} \geq n/(2d^2), \forall i \in \mathcal{I}^*, \forall j \in \mathcal{J}^*\right\}.$$

From the Hoeffding bound, we know that

$$P[\mathcal{E}'_n] \geq 1 - d^2 \cdot \exp(-n/(2d^4)).$$

We now set a N'_1 such that the above probability bound is small enough when $n \geq N'_1$. To be specific, N'_1 is set as

$$N'_1 = 2d^4 \cdot \log(4d^6/\varepsilon). \tag{84}$$

Then, denote by $\mathcal{E}'$ the event such that

$$\mathcal{E}' = \{\cap_{n \geq N'_1} \mathcal{E}'_n\}.$$

From the union bound, we have

$$P[\mathcal{E}'] \geq 1 - \sum_{n=N'_1}^{N} d^2 \cdot \exp(-n/(2d^4)) \geq 1 - 2d^6 \cdot \exp(-N'_1/(2d^4)) = 1 - \varepsilon/2.$$

Conditional on the event $\mathcal{E}'$ happens, we know that the absolute value of each element of $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}(\mathcal{H}^n) - A_{\mathcal{I}^*,\mathcal{J}^*}$ is upper bounded by $\mathrm{Rad}(n/(2d^2), \varepsilon/d^2)$, given that the following event

$$\mathcal{E} = \left\{\left|\frac{1}{n_{i,j}} \cdot \sum_{q=1}^{n_{i,j}} A_q(i,j) \cdot \mathbb{I}\{i_q = i, j_q = j\} - A_{i,j}\right| \leq \mathrm{Rad}(n/(2d^2), \varepsilon/(2d^2)), \ \ \forall i \in \mathcal{I}^*, j \in \mathcal{J}^*\right\} \tag{85}$$

is assumed to happen (it holds with probability at least $1 - \varepsilon/2$ following standard Hoeffding bound). Therefore, from the union bound, we know that both event $\mathcal{E}'$ and event $\mathcal{E}$ happen with probability at least $1 - \varepsilon$.

Conditional on the events $\mathcal{E}'$ and $\mathcal{E}$ happen, for $n \geq \max\{N'_0, N'_1\} = N'_0$, it holds that

$$|\mathbb{E}_{i_n,j_n}[\xi_j(n)|\mathcal{H}^n]| \leq \frac{1}{N-n} \cdot \left\|\left(\widehat{A}_{\mathcal{I}^*,j}(\mathcal{H}^n) - A_{\mathcal{I}^*,j}\right) \cdot \mathbf{x}^n_{\mathcal{I}^*}\right\|_\infty \leq \frac{\mathrm{Rad}(n/(2d^2), \varepsilon/(2d^2))}{N-n}. \tag{86}$$

Then, we know that

$$\begin{aligned}
\frac{\sum_{n=N'_0}^{N'-1} \left|\mathbb{E}[\xi^n_j|\mathcal{H}^n]\right|}{\sqrt{2} \cdot d} &\leq \sqrt{\frac{\log(4d^2/\varepsilon)}{2}} \cdot \sum_{n=N'_0}^{N'-1} \frac{1}{\sqrt{n} \cdot (N-n)} \\
&\leq \sqrt{\frac{\log(4d^2/\varepsilon)}{2}} \cdot \sqrt{N'-1} \cdot \sum_{n=N'_0}^{N'-1} \frac{1}{n \cdot (N-n)} \\
&= \sqrt{\frac{\log(4d^2/\varepsilon)}{2}} \cdot \frac{\sqrt{N'-1}}{N} \cdot \sum_{n=N'_0}^{N'-1} \left(\frac{1}{n} + \frac{1}{N-n}\right) \\
&\leq \sqrt{2\log(4d^2/\varepsilon)} \cdot \frac{\sqrt{N'-1}}{N} \cdot \log(N) \leq \frac{\sqrt{2\log(4d^2/\varepsilon)}}{\sqrt{N}} \cdot \log(N) \\
&\leq \frac{\eta}{2}
\end{aligned} \tag{87}$$

for a N large enough such that

$$N \geq \frac{16\log(4d^2/\varepsilon)}{\eta^2} = \frac{1024d^3 \cdot \|A_{\mathcal{I}^*,\mathcal{J}^*}\|_2^2 \cdot \log(4d^2/\varepsilon)}{\sigma^2}. \tag{88}$$

Combining Equation (87) and Equation (83) with $b = \eta/2$, and apply a union bound over all $j \in \mathcal{J}^*$, as well as event $\mathcal{E}$ and $\mathcal{E}'$, we know that

$$P[\tau \leq N'] \leq \frac{\eta^2}{4d^4} \cdot \exp\left(-\frac{\eta^2 \cdot (N-N')}{8d^4}\right) + \varepsilon. \tag{89}$$

Therefore, we know that (we set $\varepsilon = 1/N$)

$$\mathbb{E}[N-\tau] = \sum_{N'=1}^{N} P[\tau \leq N'] \leq N'_0 + 1 + \sum_{N'=\bar{N}_0}^{N} P[\tau \leq N'] \leq N'_0 + 1 + 2 \cdot \exp(-\eta^2/(8d^4))$$

which completes our proof.

J Proof of Theorem 8

We now present the worst-case bound on the suboptimality of Algorithm 2 and Algorithm 3 with a finite sample size N. Denote by $\hat{\mathcal{I}}$ and $\hat{\mathcal{J}}$ the output of Algorithm 2, we know from Theorem 6 that with a probability at least $1-\varepsilon$, the solution $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$ forms a feasible solution to V^{Prime} (by properly defining variable μ) with an sub-optimality gap bounded by

$$\text{Rad}(N/m, \varepsilon/m) + 4d^2\widehat{\kappa} \cdot \text{Rad}(N/m, \varepsilon/m),$$

where $\hat{\sigma}_{\min}$ and $\hat{\sigma}_{\max}$ are the smallest and the largest absolute value of the singular values of the matrix

$$\begin{bmatrix} \widehat{A}^\top_{\hat{\mathcal{I}},\hat{\mathcal{J}}} & -\mathbf{e}^{|\hat{\mathcal{J}}|} \\ (\mathbf{e}^{|\hat{\mathcal{I}}|})^\top & 0 \end{bmatrix}.$$

We now further denote by $\bar{\mathbf{x}}$ the output of Algorithm 3 and we compare $\bar{\mathbf{x}}$ with $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$. Note that the procedure of Algorithm 3 is used to approximate the solution of the linear equations

$$\begin{bmatrix} A^\top_{\hat{\mathcal{I}},\hat{\mathcal{J}}} & -\mathbf{e}^{|\hat{\mathcal{J}}|} \\ (\mathbf{e}^{|\hat{\mathcal{I}}|})^\top & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}^*_{\hat{\mathcal{I}}}(\hat{\mathcal{I}}, \hat{\mathcal{J}}) \\ \mu^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{90}$$

which is exactly $\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}})$. Further note that from Theorem 7, we have that (we remove the dependency on the sample size N_0 and follows from (70) by setting $\varepsilon = 1/N$)

$$\left\|\mathbf{x}^*(\hat{\mathcal{I}}, \hat{\mathcal{J}}) - \mathbb{E}[\bar{\mathbf{x}}]\right\|_2 \le \mathcal{O}\left(\sqrt{d}\kappa(A^*) \cdot \left(\frac{d^5}{\sigma^2}+1\right) \cdot \log(N) \cdot \frac{1}{N}\right). \tag{91}$$

Therefore, we know that $\mathbb{E}[\bar{\mathbf{x}}]$ is a feasible solution to V^{Prime} with an optimality gap bounded by (we set $\varepsilon = 1/N$)

$$\mathcal{O}\left(\sqrt{d}\kappa(A^*) \cdot \left(\frac{d^5}{\sigma^2}+1\right) \cdot \frac{\log(N)}{N} + d^2\hat{\kappa} \cdot \sqrt{\frac{m\log(N)}{N}}\right). \tag{92}$$

Our proof is thus completed.

K Proof of Theorem 9

We now compare the instance-dependent bound presented in Theorem 7 with the worst-case bound presented in Theorem 8. On one hand, from Theorem 8, we know that $\mathbb{E}[\bar{\mathbf{x}}]$ is a feasible solution to V^{Prime} with an optimality gap bounded by

$$\mathcal{O}\left(\sqrt{d}\kappa(A^*) \cdot \left(\frac{d^5}{\sigma^2}+1\right) \cdot \frac{\log(N)}{N} + d^2\hat{\kappa} \cdot \sqrt{\frac{m\log(N)}{N}}\right). \tag{93}$$

On the other hand, from the instance-dependent bound in Theorem 7 and noting

$$\|A^\top\mathbf{x}^* - A^\top\mathbb{E}[\bar{\mathbf{x}}]\|_\infty \le \|A^\top\mathbf{x}^* - A^\top\mathbb{E}[\bar{\mathbf{x}}]\|_2 \le \|A\|_2 \cdot \|\mathbf{x}^* - \mathbb{E}[\bar{\mathbf{x}}]\|_2 \le d \cdot \|\mathbf{x}^* - \mathbb{E}[\bar{\mathbf{x}}]\|_2$$

we know that $\mathbb{E}[\bar{\mathbf{x}}]$ is a feasible solution to V^{Prime} with an optimality gap bounded by (follows from (70) by setting $\varepsilon = 1/N$)

$$\mathcal{O}\left(d^{3/2}\kappa(A^*) \cdot \left(\frac{d^5}{\sigma^2}+1\right) \cdot \frac{\log(N)}{N}\right), \tag{94}$$

as long as

$$N \ge \mathcal{O}\left(\frac{d^5}{\sigma^2} + \frac{m}{\delta^2}\right) \cdot \log(2mN). \tag{95}$$

Then, we translate the above bounds into the sample complexity bound. Note that the bound (93) implies that in order for $\mathbb{E}[\bar{\mathbf{x}}]$ to be a feasible solution to V^{Prime} with an optimality gap bounded by ε, the number of samples N needs to satisfy the condition (where we upper bound $\kappa(A^*)$ by d/σ)

$$N = \mathcal{O}\left(\frac{d^{13/2}}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + \frac{d^4 \cdot \hat{\kappa}^2 \cdot m}{\varepsilon^2}\right). \tag{96}$$

Moreover, the bounds in (94) and (95) imply that in order for $\mathbb{E}[\bar{\mathbf{x}}]$ to be a feasible solution to V^{Prime} with an optimality gap bounded by ε, the number of samples N only needs to satisfy the condition

$$N = \mathcal{O}\left(\left(\frac{d^{15/2}}{\sigma^3} + \frac{m}{\delta^2}\right) \cdot \frac{\log(m/\varepsilon)}{\varepsilon}\right) \tag{97}$$

We classify into the following two scenarios:
i) the instance-dependent bound is smaller. Note that when the bound in Equation (97) is smaller than the bound in Equation (96), we have that

$$\frac{1}{\delta^2} \leq \mathcal{O}\left(\frac{d^4 \cdot \hat{\kappa}^2}{\varepsilon}\right)$$

In this case, as long as the sample size N satisfies the condition in (97), the condition in (95) is also satisfied, and therefore the instance-dependent guarantee is achievable (implied by Algorithm 2 finds the optimal basis with a high probability). Then, in this case, we know that a ε-suboptimality gap NE can be obtained by our Algorithm 2 and Algorithm 3 with a sample complexity bound of

$$\begin{aligned}
&\mathcal{O}\left(\left(\frac{d^{15/2}}{\sigma^3} + \frac{m}{\delta^2}\right) \cdot \frac{\log(m/\varepsilon)}{\varepsilon}\right) \\
=&\mathcal{O}\left(\frac{d^{15/2}}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + m \cdot \min\left\{\frac{d^4 \cdot \widehat{\kappa}^2}{\varepsilon^2}, \frac{1}{\delta^2 \cdot \varepsilon}\right\} \cdot \log(m/\varepsilon)\right).
\end{aligned}$$

ii) the instance-independent bound is smaller. Note that when the bound in Equation (96) is smaller than the bound in Equation (97), we have that

$$\frac{1}{\delta^2} \geq \mathcal{O}\left(\frac{d^4 \cdot \hat{\kappa}^2}{\varepsilon}\right)$$

In this case, we can directly apply the instance-independent bound and we know that a ε-suboptimality gap NE can be obtained by our Algorithm 2 and Algorithm 3 with a sample complexity bound of

$$\begin{aligned}
&\mathcal{O}\left(\frac{d^{15/2}}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + \frac{d^4 \cdot \hat{\kappa}^2 \cdot m}{\varepsilon^2}\right) \\
=&\mathcal{O}\left(\frac{d^{15/2}}{\sigma^3} \cdot \frac{\log(m/\varepsilon)}{\varepsilon} + m \cdot \min\left\{\frac{d^4 \cdot \widehat{\kappa}^2}{\varepsilon^2}, \frac{1}{\delta^2 \cdot \varepsilon}\right\} \cdot \log(m/\varepsilon)\right).
\end{aligned}$$

Our proof is thus completed.

L Parameters Estimation

We first show how to estimate some unknown problem parameters. We show that the gap δ in Definition 5 can be estimated efficiently with finite sample. Here, we assume that there exists an oracle algorithm that can solve the following minimum non-zero suboptimality gap problem

$$\min_{\mathcal{I} \subset [m_0]} \{V - V_{\mathcal{I}} : V - V_{\mathcal{I}} > 0\} \tag{98}$$

for a given index set $[m_0]$ and a given LP V. Our algorithm is formally presented in Algorithm 4. Note that we iteratively sample $i_n \in [m_1]$ and $j_n \in [m_2]$ lexicographically such that each matrix entry $A_{i,j}$ will be sampled sufficiently many times.

We provide a brief explanation on why Algorithm 4 can be used to give a lower bound of the gap δ. Note that following standard Hoeffding inequality, for any $\varepsilon > 0$, we know that the gap between each element of the matrix A and $\widehat{A}$ is upper bounded by $\text{Rad}(n/m, \varepsilon/m)$ with probability at least $1 - \varepsilon$. Under this high probability event, we know that the value between $V_{\mathcal{I}}^{\text{Prime}}$ and $\widehat{V}_{\mathcal{I}}^{\text{Prime}}$, as well as $V_{\mathcal{I},\mathcal{J}}^{\text{Dual}}$ and $\widehat{V}_{\mathcal{I},\mathcal{J}}^{\text{Dual}}$, for any $\mathcal{I} \subset [m_1]$ and $\mathcal{J} \subset [m_2]$, is upper bounded by $\text{Rad}(n/m, \varepsilon/m)$. This result further implies that

$$\Pr\left[|\delta_1 - \delta_1'| \geq 2 \cdot \text{Rad}(n/m, \varepsilon/m)\right] \leq \varepsilon \text{ and } \Pr\left[|\delta_2 - \delta_2'| \geq 2 \cdot \text{Rad}(n/m, \varepsilon/m)\right] \leq \varepsilon.$$

Then, it holds that

$$\Pr\left[|\delta - \delta'| \geq 2 \cdot \text{Rad}(n/m, \varepsilon/m)\right] \leq \varepsilon.$$

Therefore, as long as $\delta' \geq 4 \cdot \text{Rad}(n/m, \varepsilon/m)$, we know that $\frac{\delta'}{2} \leq \delta$ with a probability at least $1 - \varepsilon$. Our result is formalized as follows.

Algorithm 4 Algorithm to lower bound δ

Input: a failure probability $\epsilon > 0$, an oracle algorithm that can solve the problem in (98).
for $n = 1, 2, \dots,$ until stopped, **do**
Let $a_n = (n \bmod m) + 1$ and compute $a_n = j_n \cdot m_1 + i_n$ with i_n and j_n are integers and $i_n \in \{1, \dots, m_1\}$ and $j_n \in \{1, \dots, m_2\}$.
Take the primal action i_n and the dual action j_n, and observe $\widetilde{A}_{i_n,j_n} = A_{i_n,j_n} + \eta_n$.
Update $\mathcal{H}^{n+1} = \mathcal{H}^n \cup \{(i_n, j_n, \widetilde{A}_{i_n,j_n}\}$.
Construct $\widehat{A}$ and the LP $\widehat{V}^{\text{Prime}}$.
Call the oracle algorithm to compute

$$\delta_1' = \min_{\mathcal{I} \subset [m_1]} \{\widehat{V}^{\text{Prime}} - \widehat{V}_{\mathcal{I}}^{\text{Prime}} : \widehat{V}^{\text{Prime}} - \widehat{V}_{\mathcal{I}}^{\text{Prime}} > 0\} \tag{99}$$

and compute

$$\delta_2' = \min_{\mathcal{I} \subset [m_1], \mathcal{J} \subset [m_2]} \{\widehat{V}^{\text{Dual}} - \widehat{V}_{\mathcal{I},\mathcal{J}}^{\text{Dual}} : \widehat{V}^{\text{Dual}} - \widehat{V}_{\mathcal{I},\mathcal{J}}^{\text{Dual}} > 0\}. \tag{100}$$

Stop if $\delta' = \min\{\delta_1', \delta_2'\} \geq 4 \cdot \text{Rad}(n/m, \varepsilon/m)$.
Output: the estimation δ'.

Theorem 10 *Denote by δ' the output of Algorithm 4 and denote by N_1 the number of samples used in Algorithm 4. Then, with a probability at least $1 - \varepsilon$, it holds that*

$$\frac{\delta'}{2} \leq \delta \leq 2\delta' \ \text{ and } \ N_1 \leq \frac{18m}{\delta^2} \cdot \log(2m/\varepsilon).$$

Though the bound on N_1 in Theorem 10 depends on the true value of δ, however, our Algorithm 4 does not require the knowledge of δ.

Next, we show that the parameter σ, the minimum singular value of A^* defined in Equation (23), can be efficiently estimated using finite samples. The algorithm is summarized in Algorithm 5. We provide a brief explanation on why Algorithm 5 can be used to give a good estimation of the parameter σ. Note that following standard Hoeffding inequality, for any $\varepsilon > 0$, we know that the absolute value of the matrix $\Delta A = A_{\mathcal{I}^*,\mathcal{J}^*} - \widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}$ is upper bounded by $\text{Rad}(n/d^2, \varepsilon/d^2)$ with a probability at least $1 - \varepsilon$. Then following Theorem 1 of [43], we know that the difference of the corresponding singular values of $A_{\mathcal{I}^*,\mathcal{J}^*}$ and $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}$ (sorted from the largest to the smallest) is upper bounded by $\|\Delta A\|_2$, which is again upper bounded by $d \cdot \text{Rad}(n/d^2, \varepsilon/d^2)$ with a probability at least $1 - \varepsilon$. Formally, it holds that

$$\Pr\left[|\sigma - \sigma'| \geq d \cdot \text{Rad}(n/d^2, \varepsilon/d^2)\right] \leq \varepsilon. \tag{101}$$

Therefore, as long as $\sigma' \geq 2d \cdot \text{Rad}(n/d^2, \varepsilon/d^2)$, we know that $\frac{\sigma'}{2} \leq \sigma$ with a probability at least $1 - \varepsilon$.

Algorithm 5 Algorithm to estimate σ

Input: a failure probability $\epsilon > 0$, the support set $\mathcal{I}^*$ and $\mathcal{J}^*$.
for $n = 1, 2, \dots,$, **do**
Let $a_n = (n \bmod d^2) + 1$ and compute $a_n = b_n \cdot d + c_n$ with b_n and c_n are integers and $b_n \in \{1, \dots, d\}$ and $c_n \in \{1, \dots, d\}$.
Iteratively sample a pair $(i_n, j_n) \sim \mathcal{I}^* \times \mathcal{J}^*$ such that i_n is the b_n-th element in $\mathcal{I}^*$ and j_n is the c_n-th element in $\mathcal{J}^*$.
Take the primal action i_n and the dual action j_n, and observe $\widetilde{A}_{i_n,j_n} = A_{i_n,j_n} + \eta_n$.
Update $\mathcal{H}^{n+1} = \mathcal{H}^n \cup \{(i_n, j_n, \widetilde{A}_{i_n,j_n}\}$.
Construct $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}(\mathcal{H}^{n+1})$ and let σ' denote the smallest absolute value of its singular values.
Stop if $\sigma' \geq 2d \cdot \text{Rad}(n/d^2, \varepsilon/d^2)$.
Output: the estimation σ'.

Theorem 11 *Denote by σ' the output of Algorithm 5 and denote by N_2 the number of samples used in Algorithm 5. Then, with a probability at least $1-\varepsilon$, we have $\frac{\sigma'}{2} \le \sigma \le 2\sigma'$ and it holds that*

$$\Pr\left[N_2 \le \frac{9d^4}{2\sigma^2} \cdot \log(2d^2/\varepsilon)\right] \ge 1-\varepsilon.$$

Note that Theorem 11 ensures that $\frac{\sigma'}{2}$ is a lower bound on σ and we also obtain a high probability bound on the number of samples needed.

L.1 Proof of Theorem 10

Note that following standard Hoeffding inequality, for any $\varepsilon > 0$, we know that the gap between each element of the matrix A and $\widehat{A}$ is upper bounded by $\mathrm{Rad}(n/m, \varepsilon/m)$ with a probability at least $1-\varepsilon$. Then, from Claim 2 and Claim 3, we know that the value between $V_{\mathcal{I}}^{\mathrm{Prime}}$ and $\widehat{V}^{\mathrm{Prime}}$, as well as $V_{\mathcal{I},\mathcal{J}}^{\mathrm{Dual}}$ and $\widehat{V}_{\mathcal{I},\mathcal{J}}$, for any $\mathcal{I} \subset [m_1]$ and $\mathcal{J} \subset [m_2]$, is upper bounded by $\mathrm{Rad}(n/m, \varepsilon/m)$, with a probability at least $1-\varepsilon$. This result further implies that

$$\Pr\left[|\delta_1 - \delta_1'| \ge 2 \cdot \mathrm{Rad}(n/m, \varepsilon/m)\right] \le \varepsilon \ \text{ and } \ \Pr\left[|\delta_2 - \delta_2'| \ge 2 \cdot \mathrm{Rad}(n/m, \varepsilon/m)\right] \le \varepsilon.$$

Then, it holds that

$$\Pr\left[|\delta - \delta'| \ge 2 \cdot \mathrm{Rad}(n/m, \varepsilon/m)\right] \le \varepsilon.$$

Therefore, as long as $\delta' \ge 4 \cdot \mathrm{Rad}(n/m, \varepsilon/m)$, we know that $\frac{\delta'}{2} \le \delta \le 2\delta'$ with a probability at least $1-\varepsilon$. On the other hand, in order to bound N_1, we know that with a probability at least $1-\varepsilon$, it holds that

$$\delta \le \delta' + 2 \cdot \mathrm{Rad}(N_1/m, \varepsilon/m),$$

which implies that

$$N_1 \le \frac{18m}{\delta^2} \cdot \log(2m/\varepsilon).$$

Our proof is thus completed.

L.2 Proof of Theorem 11

Note that following standard Hoeffding inequality, for any $\varepsilon > 0$, we know that the absolute value of the matrix $\Delta A = A_{\mathcal{I}^*,\mathcal{J}^*} - \widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}$ is upper bounded by $\mathrm{Rad}(n/d^2, \varepsilon/d^2)$ with a probability at least $1-\varepsilon$. Then following Theorem 1 of [43], we know that the difference of the corresponding singular values of $A_{\mathcal{I}^*,\mathcal{J}^*}$ and $\widehat{A}_{\mathcal{I}^*,\mathcal{J}^*}$ (sorted from the largest to the smallest) is upper bounded by $\|\Delta A\|_2$, which is again upper bounded by $d \cdot \mathrm{Rad}(n/d^2, \varepsilon/d^2)$ with a probability at least $1-\varepsilon$. Formally, it holds that

$$\Pr\left[|\sigma - \sigma'| \ge d \cdot \mathrm{Rad}(n/d^2, \varepsilon/d^2)\right] \le \varepsilon.$$

Therefore, as long as $\sigma' \ge 2d \cdot \mathrm{Rad}(n/d^2, \varepsilon/d^2)$, we know that $\frac{\sigma'}{2} \le \sigma \le 2\sigma'$ with a probability at least $1-\varepsilon$. On the other hand, in order to bound N_2, we know that with a probability at least $1-\varepsilon$, we have

$$\sigma \le \sigma' + d \cdot \mathrm{Rad}(N_2/d^2, \varepsilon/d^2),$$

which implies that

$$N_2 \le \frac{9d^4}{2\sigma^2} \cdot \log(2d^2/\varepsilon).$$

Our proof is thus completed.