MYERSON-GUIDED MACHINE LEARNING FOR NEAR-OPTIMAL AUCTION DESIGN

Anonymous authors
Paper under double-blind review

ABSTRACT

Myerson characterizes the revenue-optimal auction for single-item settings, but the optimal mechanism remains unknown in general multi-bidder, multi-item environments. Neural mechanisms provide an expressive approach to this problem, yet some existing methods achieve only limited revenue gains over Item-wise Myerson in larger auction settings. Our profile-level analysis of ALGNet reveals complementary strengths: its gains on low-competition profiles are consistent with cross-item interactions and probabilistic allocations, but it increasingly underperforms Item-wise Myerson as item-level competition intensifies. Motivated by this observation, we propose Myerson-guided Machine Learning (M2L), a three-stage framework that uses Item-wise Myerson as a structural teacher. M2L first imitates Myerson's allocation and payment rules, then autonomously explores cross-item interactions and probabilistic allocations, and finally targets profiles on which the learned mechanism still underperforms Myerson while preserving its advantages elsewhere. Experiments across multi-item benchmarks show that M2L is competitive with, and often outperforms, existing neural approaches in expected revenue, substantially reduces below-Myerson profiles, and maintains low empirically estimated ex-post regret.

1 INTRODUCTION

A central goal in mechanism design is to construct auctions that maximize the seller's expected revenue while satisfying dominant-strategy incentive compatibility (DSIC) and individual rationality (IR). In single-item settings, Myerson's seminal work (Myerson, 1981) characterizes the revenue-optimal mechanism and provides a foundation for optimal auction design. In general multi-bidder, multi-item settings, the optimal mechanism remains unknown despite decades of research (Alaei et al., 2012; Babaioff et al., 2020; Cai & Zhao, 2017; Hart & Nisan, 2017; Pavlov, 2011; Wang & Tang, 2014; Yao, 2017).

The central challenge is the theoretical intractability of auction design with multiple bidders and items. This difficulty has motivated a growing body of work that uses deep learning to automatically generate auction mechanisms. Existing neural approaches can be broadly divided into two paradigms.

The first is menu-based neural mechanism design, including AMenuNet (Duan et al., 2023) and GemNet (Wang et al., 2024). These methods construct menus from which bidders select their preferred options, thereby satisfying DSIC by design. However, this guarantee comes at the cost of a restricted mechanism space. A fixed-size menu limits the expressiveness of the learned mechanism, and the resulting mechanism class may not contain the optimal auction (Babaioff et al., 2017; Hart & Nisan, 2019). These methods also face scalability challenges: the required menu complexity grows rapidly with the number of bidders and items, leading to higher computational costs.

The second paradigm directly parameterizes allocation and payment rules using neural networks, such as RegretNet (Dütting et al., 2019), ALGNet (Rahme et al., 2021b), and RegretFormer (Ivanov et al., 2022). These methods search over a substantially richer class of mechanisms and therefore have greater potential to represent high-revenue solutions. However, the increased expressiveness comes with a weaker incentive guarantee: DSIC is typically pursued through ex-post regret minimization, and the resulting mechanisms may consequently exhibit small DSIC violations. In this work, we follow this direct-parameterization paradigm because its expressive mechanism space provides a promising foundation for learning high-revenue auctions while explicitly controlling regret.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Despite their expressive power, direct-parameterization methods are evaluated primarily through expected revenue and regret, which provide little insight into the mechanisms learned by these black-box neural networks. Existing structural analyses are largely limited to small settings such as 2 bidders and 2 items, where allocation rules can be visualized directly. For larger auctions, the profile-level behavior underlying the reported aggregate performance remains poorly understood.

To gain more insight, we use Item-wise Myerson as a reference. As the optimal single-item auction, Item-wise Myerson offers a simple, interpretable reference for analyzing learned multi-item mechanisms. We compare ALGNet and Item-wise Myerson on the same valuation profiles in auctions with 10 items and 2–7 additive bidders. We use $n \times m$ to denote a setting with n bidders and m items. We find that the fraction of profiles on which ALGNet earns less revenue than Item-wise Myerson increases from approximately 11% in the 2×10 setting to 71% in the 7×10 setting, eventually causing ALGNet's average revenue to fall below Item-wise Myerson.

We further identify a systematic relationship between these below-Myerson cases and item-level competition. We define an item as highly competitive if its second-highest valuation exceeds the Myerson reserve. For each valuation profile, we let S denote the number of highly competitive items. As S increases, the fraction of profiles on which ALGNet underperforms Item-wise Myerson also increases. Moreover, profiles with large S become more prevalent as the number of bidders increases, explaining ALGNet's declining average performance relative to Myerson.

Our experiments reveal a key limitation of existing direct-parameterization methods. On low-competition profiles, their revenue advantage is consistent with their ability to use cross-item interactions and probabilistic allocations, including on items that Item-wise Myerson would leave unsold. When a profile contains many highly competitive items, however, Item-wise Myerson already extracts substantial revenue through item-level competition, and cross-item interactions and probabilistic allocations can weaken this advantage and reduce revenue. To combine these complementary strengths, we propose Myerson-guided Machine Learning (M2L), a three-stage framework that uses Item-wise Myerson as a structured teacher. M2L is designed to exploit cross-item interactions and probabilistic allocations on low-competition profiles while preserving, as much as possible, Myerson's revenue advantage when many items are highly competitive. Stage I explicitly imitates Myerson's allocation and payment rules to acquire an auction-theoretic structural prior. Stage II then departs from this prior through autonomous revenue–regret optimization, allowing the mechanism to discover cross-item interactions and probabilistic allocations. Stage III does not imitate Myerson again; instead, it selectively changes the forward optimization state on profiles where the explored mechanism underperforms Myerson. M2L thus follows a *learn the prior, depart from the prior, and selectively repair harmful departures* progression.

Experiments show that M2L achieves the highest revenue among direct-parameterization methods while maintaining low regret. M2L also outperforms most existing neural baselines. Compared with OD-VVCA, M2L achieves higher revenue in smaller settings and remains competitive in settings with more bidders and items. Profile-level results show that Stage II reduces below-Myerson cases, although a substantial fraction remains as competition increases. Stage III substantially reduces these remaining cases and rarely lowers profile-level revenue relative to Stage II, thereby improving performance relative to Myerson while preserving the gains from autonomous exploration.

2 Related work

RegretNet (Dütting et al., 2019) established the foundational framework by maximizing revenue while treating dominant-strategy incentive compatibility (DSIC) as a soft constraint. Ivanov et al. (2022) subsequently enhanced this architecture by incorporating attention mechanisms to better capture bidder interactions. Rahme et al. (2021a) introduced permutation-equivariant networks that enforce symmetry in the mechanism. To address training stability, ALGNet (Rahme et al., 2021b) reformulated training as a min-max game between the auctioneer and a misreporting agent. Building on this adversarial approach, You et al. (2023) introduced discrete distribution sampling and a fine-grained payment module.

Conversely, parallel research ensures exact strategyproofness by restricting learning to theoretically valid classes of mechanisms. Curry et al. (2022) proposed a differentiable framework to learn the parameters of Affine Maximizer Auctions (AMAs). To overcome the computational bottlenecks

of AMAs, Duan et al. (2023) developed a neural network that efficiently co-optimizes allocation menus. In a different structural direction, Shen et al. (2019) designed a menu-based architecture that, by construction, guarantees strategyproofness in single-bidder settings. Most recently, Wang et al. (2024) extended this menu-based paradigm to multi-bidder environments via GemNet. However, these strictly constrained methods often limit the solution space, resulting in lower revenue than unconstrained approximations.

Beyond standard revenue maximization, the deep learning framework has been successfully adapted to accommodate diverse economic constraints and settings. Feng et al. (2018) addressed the challenge of private budget constraints, while Kuo et al. (2020) introduced fairness considerations into the allocation rules. To align mechanisms with human values, Peri et al. (2021) incorporated preference learning. Further extensions include the integration of contextual information by Duan et al. (2022) and, recently, the handling of correlated valuations by Huo et al. (2025).

3 PRELIMINARIES

We consider auctions where n bidders, represented by the set $N = \{1, \ldots, n\}$, try to outbid each other for m items, represented by the set $M = \{1, \ldots, m\}$. Each bidder i has a valuation function $v_i : 2^M \to \mathbb{R}_{\geq 0}$. We focus on *additive* valuation functions, that is, a bidder i's valuation for the set $S \subseteq M$ of items is the sum of the valuations for each item in S. Formally, $v_i(S) = \sum_{j \in S} v_i(j)$ for all $S \subseteq M$, where v_i is independently drawn from a distribution F_i over possible valuation functions V_i. We denote the set of all possible valuation profiles as $V = \prod_{i=1}^{n} V_i$, and $v = (v_1, \ldots, v_n) \in V$ is the valuation profile.

Although the auctioneer knows the distributions $F = (F_1, \ldots, F_n)$, the actual valuations v of the bidders are unknown. Bidders report their bids, which may not be truthful, and an auction mechanism determines the allocation of items to the bidders and the corresponding payments. An auction mechanism (a, p) consists of allocation rules $a_i : V \to [0, 1]^m$ and payment rules $p_i : V \to \mathbb{R}_{\geq 0}$, where $a_{ij}(b)$ is the marginal probability that bidder i receives item j under bid profile b. Feasibility requires $\sum_{i=1}^{n} a_{ij}(b) \leq 1$ for every item j and bid profile b. Given $b = (b_1, \ldots, b_n) \in V$, the auction computes the allocation probabilities $a(b)$ and payments $p(b)$. We denote the valuation profile v without bidder i as v_{-i}, and similarly for bid profile b_{-i}. The set of possible valuation profiles for all bidders except bidder i is denoted as $V_{-i} = \prod_{j \neq i} V_j$.

A bidder's expected utility, where the expectation is over the mechanism's internal randomization, is

$$u_i(v_i, b) = \sum_{j=1}^{m} v_{ij} a_{ij}(b) - p_i(b),$$

where $v_{ij} = v_i(j)$. An auction is dominant-strategy incentive compatible (DSIC) if truthful reporting is the optimal strategy for every bidder, regardless of the other bidders' bids. In formal terms, this means that $u_i(v_i, (v_i, b_{-i})) \geq u_i(v_i, (b_i, b_{-i}))$ for all bidders i, all valuations $v_i \in V_i$, all bids $b_i \in V_i$, and all bids $b_{-i} \in V_{-i}$ from others. On the other hand, an auction is said to be individually rational (IR) if every bidder obtains nonnegative expected utility when reporting truthfully, meaning that $u_i(v_i, (v_i, b_{-i})) \geq 0$.

A DSIC auction is designed so that each bidder has a dominant strategy to report the true valuation, and the revenue on the valuation profile v is $\sum_i p_i(v)$. The goal is to identify an auction mechanism that satisfies DSIC and IR and maximizes the revenue. For DSIC, we use the ex-post regret to measure the violation of DSIC, defined as follows:

$$Rgt_i = \max_{v_i' \in V_i} u_i\left(v_i, (v_i', v_{-i})\right) - u_i\left(v_i, (v_i, v_{-i})\right)$$

We parameterize the auction mechanism (a, p) using neural networks (a_θ, p_θ) with parameters θ. As exact DSIC cannot be strictly enforced, we use ex-post regret to quantify the violation of DSIC. We reformulate optimal auction design as a constrained learning problem, where the objective is to maximize expected revenue while minimizing ex-post regret over sampled valuation profiles.

Table 1: Profile-level comparison between ALGNet and Item-wise Myerson on $n \times 10$ settings. The gap is defined as ALGNet revenue minus item-wise Myerson revenue.

Setting	Myerson Rev.	ALGNet Rev.	Gap	Below-Myerson Profiles
2×10	4.1641	4.8781	$+0.7140$	11.40%
3×10	5.3137	5.5669	$+0.2533$	29.08%
4×10	6.1235	6.3225	$+0.1989$	31.18%
5×10	6.7213	6.7812	$+0.0599$	45.78%
6×10	7.1656	7.1044	-0.0612	63.74%
7×10	7.5064	7.4032	-0.1032	71.32%

4 Profile-Level Diagnosis of ALGNet

To understand where a learned mechanism gains or loses revenue, we conduct a profile-level diagnostic study of ALGNet against Item-wise Myerson, a simple and interpretable reference that applies the optimal single-item auction independently to each item. We consider additive auctions with 10 items and $n = 2, \ldots, 7$ bidders, with valuations sampled independently from $U[0, 1]$, and identify the profiles on which ALGNet outperforms or underperforms this baseline.

4.1 Profile-Level Comparison

For each setting, we sample 100,000 valuation profiles and compare the revenue of ALGNet and Item-wise Myerson profile by profile. Table 1 reports the average revenue gap and the fraction of profiles on which ALGNet obtains lower revenue than Item-wise Myerson.

As the number of bidders increases, the fraction of profiles on which ALGNet underperforms Item-wise Myerson rises from about 11% in the 2×10 setting to about 71% in the 7×10 setting. The growing prevalence of these underperforming profiles progressively erodes ALGNet's average revenue advantage: the improvement falls below 1% in the 5×10 setting, and by the 7×10 setting, ALGNet earns lower average revenue than Item-wise Myerson.

4.2 Competition-Stratified Analysis

To determine whether ALGNet's underperformance relative to Item-wise Myerson exhibits a systematic pattern, we analyze valuation profiles by competition intensity. We define an item as highly competitive when at least two bidders value it above the reserve price. For each valuation profile, we let S denote the number of highly competitive items, with a larger S indicating stronger competition across the profile. We partition the profiles into six buckets: $S = 0$, $S = 1$–2, $S = 3$–4, $S = 5$–6, $S = 7$–8, and $S = 9$–10.

Figure 1 explains the deterioration observed in Table 1. As shown in Panel (a), the profile distribution shifts toward larger S as the number of bidders increases: low- and moderate-competition profiles are more prevalent with fewer bidders, whereas high-competition profiles increasingly dominate in larger-bidder settings.

Panels (b)–(d) reveal a systematic decline in ALGNet's relative performance as S increases. The fraction of profiles below Item-wise Myerson rises across competition buckets, while the mean revenue gap decreases from clearly positive at low or moderate S to negative at $S = 9$–10. The negative contribution to the mean gap also becomes larger in magnitude in the high-competition region. Hence, ALGNet's deterioration is concentrated on profiles with widespread item-level competition rather than being uniform across the valuation space. ALGNet's gains on lower-competition profiles are consistent with its ability to use cross-item interactions and probabilistic allocations. However, it does not reliably preserve Myerson's revenue advantage when many items are highly competitive; as these profiles become more prevalent, its overall revenue advantage diminishes.

This observation motivates using Item-wise Myerson not only as an evaluation reference but also as structural guidance during learning. Ideally, a learned mechanism should exploit cross-item interactions and probabilistic allocations when few items face strong competition while preserving the advantages of Item-wise Myerson when many items are highly competitive.

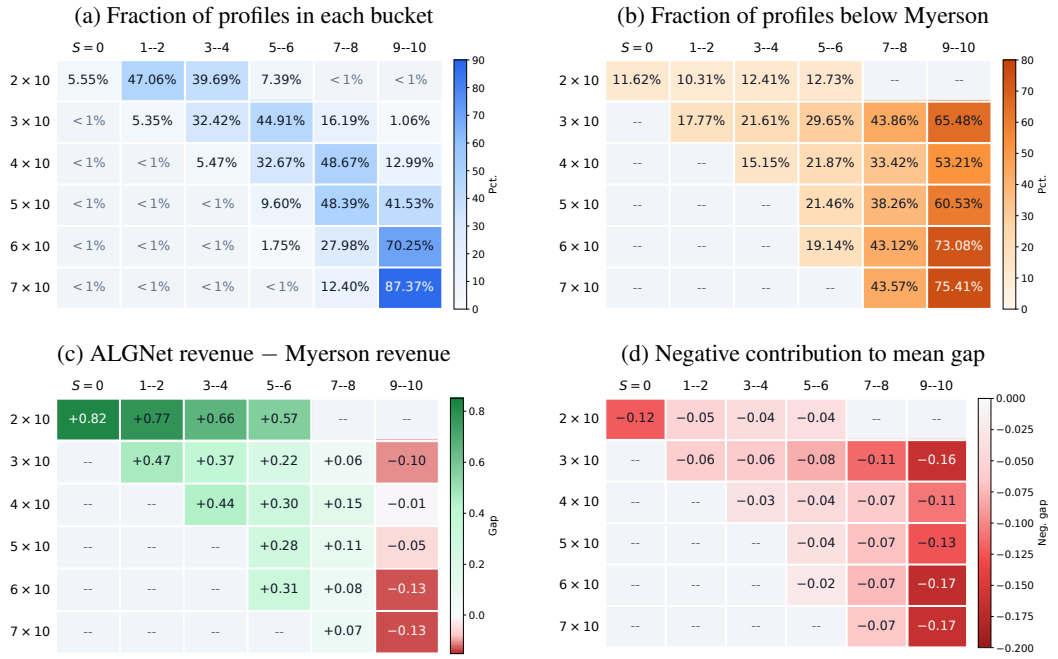

Figure 1: Competition-stratified profile-level analysis of ALGNet. Here, S is the number of items whose second-highest valuation exceeds the Myerson reserve. Panel (a) shows the distribution of valuation profiles across competition buckets. Within each bucket, Panel (b) reports the fraction of profiles on which ALGNet underperforms Item-wise Myerson, Panel (c) reports the mean revenue gap $\mathrm{Rev}_{\mathrm{ALGNet}} - \mathrm{Rev}_{\mathrm{Myerson}}$, and Panel (d) reports its negative component, $\mathbb{E}[\min\{\mathrm{Rev}_{\mathrm{ALGNet}} - \mathrm{Rev}_{\mathrm{Myerson}}, 0\}]$. Buckets containing less than 1% of the sampled profiles are marked as $< 1\%$, and their bucket-level statistics are omitted.

5 Myerson-Guided Machine Learning

In this section, we present Myerson-guided Machine Learning (M2L), a framework for learning auction mechanisms by incorporating the structural guidance of Item-wise Myerson into neural network training. Details of the neural network parameters are provided in Appendix A.

5.1 Neural Mechanism Parameterization

We consider an auction with n bidders and m items. Given a bid profile $b \in \mathbb{R}^{n\times m}$, we parameterize the mechanism using two neural networks: an allocation network a_θ and a payment network $\tilde{p}_\theta$.

The allocation network outputs

$$a_\theta(b) \in [0,1]^{n\times m},$$

which specifies the item-wise allocation probabilities. The payment network outputs an item-wise payment scaling matrix

$$\tilde{p}_\theta(b) \in [0,1]^{n\times m},$$

which specifies the fraction of the reported value charged for each allocated item. Based on these two outputs, we define the item-wise payment matrix as

$$P_\theta(b) = b \odot a_\theta(b) \odot \tilde{p}_\theta(b),$$

where $\odot$ denotes element-wise multiplication. The total payment charged to each bidder is then obtained by summing over items.

Guarantee of Individual Rationality (IR). Because each payment scaling factor $\tilde{p}_\theta(b)$ lies in $[0,1]$, a bidder's total payment cannot exceed the reported value of the resulting allocation. Under

truthful reporting, this reported value equals the bidder's true value, ensuring nonnegative utility. The mechanism therefore satisfies IR by construction.

5.2 Stage I: Structural Imitation Learning

Stage I injects the allocation and payment structure of Item-wise Myerson into the neural mechanism through supervised imitation. This equips the model with an auction-theoretic structural prior before revenue–regret optimization. Given a bid profile b sampled from the training distribution F, we apply the Item-wise Myerson auction independently to each item and obtain:

- the item-wise allocation labels
$$a_M(b) \in \{0,1\}^{n\times m},$$
- and the item-wise payment labels
$$P_M(b) \in \mathbb{R}^{n\times m}.$$

Here, $a_M(b)$ and $P_M(b)$ denote the allocation and payment outputs of the Item-wise Myerson Auction.

Using these teacher outputs, Stage I transfers the Myerson allocation and payment rules through two supervised regression objectives. The allocation loss is

$$\mathcal{L}^{\mathrm{I}}_{\mathrm{alloc}}(\theta) = \mathbb{E}_{b\sim F}\left[\|a_\theta(b) - a_M(b)\|_2^2\right].$$

The corresponding payment loss is

$$\mathcal{L}^{\mathrm{I}}_{\mathrm{pay}}(\theta) = \mathbb{E}_{b\sim F}\left[\|P_\theta(b) - P_M(b)\|_2^2\right].$$

Through supervised imitation, the neural mechanism internalizes the local competition structure encoded by Item-wise Myerson. Stage II then builds on this injected structural prior while moving beyond item-wise behavior to explore cross-item interactions and probabilistic allocations.

5.3 Stage II: Autonomous Exploration

Stage II allows the mechanism to move beyond Item-wise Myerson and explore cross-item interactions and probabilistic allocations. Following ALGNet, we formulate training as an adversarial game between the seller and a virtual misreporter, parameterized by θ and ϕ, respectively. Here, θ contains the parameters of the allocation and payment networks, while $M_\phi(v)$ represents the misreport generated for valuation profile v.

The expected revenue is evaluated under truthful reporting, whereas the virtual misreporter is used only to estimate profitable unilateral deviations in the regret term. Specifically, we define

$$P(\theta) = \mathbb{E}_{v\sim F}[\mathrm{Rev}_\theta(v)] = \mathbb{E}_{v\sim F}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij,\theta}(v)\right],$$
$$R(\theta,\phi) = \mathbb{E}_{v\sim F}\left[\mathrm{Regret}_\theta\big(v, M_\phi(v)\big)\right].$$

Thus, $P(\theta)$ depends only on the mechanism parameters because all bidders report their true valuations when revenue is computed; M_ϕ affects only the estimated regret. Here, $\mathrm{Regret}_\theta\big(v, M_\phi(v)\big) = \sum_{i=1}^{n} \mathrm{Rgt}_{i,\theta}\big(v, M_\phi(v)\big)$ is the sum of the estimated bidder-level regrets. We optimize the mechanism using the revenue–regret objective

$$\mathcal{L}_{\mathrm{II}}(\theta) = -\left(\sqrt{P(\theta)} - \sqrt{R(\theta,\phi)}\right) + R(\theta,\phi).$$

This adversarial optimization enables the mechanism to learn cross-item interactions and probabilistic allocations beyond item-wise behavior. Although Stage II improves expected revenue over Item-wise Myerson, it continues to underperform the baseline on a nontrivial fraction of valuation profiles, motivating the targeted correction introduced in Stage III.

5.4 STAGE III: MYERSON-GUIDED SELECTIVE REFINEMENT

To address these below-Myerson cases, we introduce Stage III as a Myerson-guided selective refinement stage. Item-wise Myerson serves as the teacher and the learned neural mechanism as the student. During training, their profile-level revenues determine which outputs are used in the forward pass. When Myerson obtains at least as much revenue as the student, its allocation and payment outputs are selected; otherwise, the student outputs are retained. A straight-through estimator routes the backward gradients through the student while keeping the teacher fixed. The intervention is removed at test time, and the student alone is deployed as the final mechanism.

For each bid profile b, the neural student outputs an allocation matrix $a_\theta(b)$ and an item-wise payment matrix $P_\theta(b)$, while the Item-wise Myerson teacher produces $a_M(b)$ and $P_M(b)$. Their profile-level revenues are

$$\mathcal{P}_\theta(b) = \sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij,\theta}(b), \qquad \mathcal{P}_M(b) = \sum_{i=1}^{n}\sum_{j=1}^{m} P_{ij,M}(b).$$

Based on this comparison, we define the profile-dependent intervention mask

$$m(b) = \begin{cases} 1, & \text{if } \mathcal{P}_M(b) \geq \mathcal{P}_\theta(b), \\ 0, & \text{otherwise.} \end{cases}$$

The mask determines which mechanism supplies the allocation and payment used in the training-time forward pass: $m(b) = 1$ selects the Item-wise Myerson teacher, whereas $m(b) = 0$ retains the neural student. Because this profile-dependent selection is discrete, we use a straight-through estimator to propagate gradients to the student.

Straight-through optimization. Let $z_\theta(b)$ denote the student output, $z_M(b)$ the Myerson output, and $z_I(b)$ the profile-dependent intervention output:

$$\begin{aligned} z_\theta(b) &= \big(a_\theta(b), P_\theta(b)\big), \\ z_M(b) &= \big(a_M(b), P_M(b)\big), \\ z_I(b) &= m(b) z_M(b) + (1 - m(b)) z_\theta(b). \end{aligned}$$

We define

$$\bar{z}(b) = z_\theta(b) + \mathrm{sg}\big(z_I(b) - z_\theta(b)\big),$$

where $\mathrm{sg}(\cdot)$ denotes the stop-gradient operator and $\bar{z}(b) = (\bar{a}(b), \bar{P}(b))$. The forward pass therefore evaluates the revenue–regret objective using the intervention output $z_I(b)$, while the backward pass propagates gradients through the student $z_\theta(b)$:

$$\bar{z}(b) = z_I(b), \qquad \nabla_\theta \bar{z}(b) = \nabla_\theta z_\theta(b).$$

Equivalently, for a differentiable loss L,

$$\nabla_\theta L(\bar{z}(b)) = \nabla_z L(z)\big|_{z=z_I(b)} \nabla_\theta z_\theta(b).$$

This is a straight-through surrogate gradient: the loss is evaluated at the intervention output $z_I(b)$, while its gradient is propagated through the student output $z_\theta(b)$. It is used to train the student under the intervened forward configuration. At test time, both revenue and regret are evaluated using the student mechanism alone, without intervention.

A key feature of Stage III is that the intervention affects both truthful revenue and regret. Let $\bar{p}(b) \in \mathbb{R}^n$ be the bidder-level payment vector obtained by summing $\bar{P}(b)$ over items, and let $\bar{g}_\theta = (\bar{a}, \bar{p})$ denote the resulting training-time mechanism. For a valuation profile v, bidder i's ex-post regret is

$$\mathrm{Rgt}_i(v; \bar{g}_\theta) = \max_{\hat{v}_i \in V_i} \left[u_i(v_i, (\hat{v}_i, v_{-i}); \bar{g}_\theta) - u_i(v_i, v; \bar{g}_\theta)\right],$$

where the maximization is approximated by the virtual misreporter. Because $\bar{g}_\theta$ applies profile-dependent intervention, truthful and deviating reports may activate different branches. This profile-dependent branch selection changes the forward revenue–regret landscape. The straight-through estimator then supplies a surrogate student gradient under this intervened forward configuration, without differentiating through the hard switching boundary.

Stage III retains the revenue–regret objective from Stage II but evaluates both terms using the intervention outputs $(\bar{a}(b), \bar{P}(b))$.

Under this revenue–regret objective, the intervention plays two complementary roles. On teacher-dominated profiles, the forward pass uses the Item-wise Myerson outputs, so the revenue and regret gradients are evaluated at the intervened state and propagated through the student Jacobian. When the student already outperforms Item-wise Myerson, its outputs are retained in the training-time forward pass, largely preserving such behavior learned by the neural mechanism. Stage III therefore provides targeted refinement of underperforming profiles after autonomous exploration. The intervention is training-only and deployment uses the neural mechanism alone.

6 EXPERIMENTS

In this section, we present empirical experiments to evaluate the effectiveness of M2L. All experiments are conducted on a machine equipped with an NVIDIA GeForce RTX 4090 GPU with 24 GB of memory. Results for M2L and its intermediate stages are averaged over four independent runs with distinct random seeds. To ensure a fair comparison, we evaluate M2L and all reevaluated approximately DSIC baselines using the common evaluation protocol of You et al. (2026). Final evaluation uses 10,000 test profiles. Detailed evaluation settings, network architectures, and implementation details are provided in Appendix A.

Datasets: Each valuation profile is an $n \times m$ bidder–item matrix with entries sampled i.i.d. from $U[0, 1]$. Results for an additional distribution are provided in Appendix B.

Baselines: We group the baseline methods into three categories.

(1) **Deterministic DSIC methods** include Item-wise Myerson (Myerson, 1981), which runs the optimal single-item auction independently on each item, and OD-VVCA (Duan et al., 2026).

(2) **Randomized DSIC methods** include AMenuNet (Duan et al., 2023) and GemNet (Wang et al., 2024). GemNet learns bidder-level menus but requires mixed-integer linear programming to ensure allocation feasibility, resulting in prohibitively long runtimes when the numbers of bidders and items are large and making it impractical for our large-scale settings.

(3) **Randomized approximately DSIC methods** include RegretNet (Dütting et al., 2019) and ALGNet (Rahme et al., 2021b), which control small incentive violations through regret minimization. Because RegretNet requires substantial hyperparameter tuning, we evaluate it only in the original paper's settings; ALGNet permits simpler tuning and is evaluated more broadly. We omit RegretFormer (Ivanov et al., 2022) because a subsequent evaluation reports regret as high as 0.3658 in the 3×10 setting (You et al., 2026).

Since GemNet has not released source code, we use the results reported in its original paper. For all other baselines with publicly available implementations, we rerun the released source code and evaluate the resulting mechanisms under our common evaluation protocol.

6.1 MAIN RESULTS

Our main experimental results are reported in Table 2. Among direct-parameterization methods, M2L achieves the highest reported revenue in every setting while maintaining low regret, establishing a new state of the art within this family. Compared with methods based on other mechanism classes, M2L achieves the highest overall revenue in settings with fewer bidders and items, outperforming both GemNet and OD-VVCA. In larger settings, M2L remains competitive with OD-VVCA while substantially outperforming the other baselines. These results are obtained despite the broader search space of M2L: OD-VVCA is restricted to a deterministic subclass of affine maximizer auctions, whereas M2L can represent both deterministic and randomized mechanisms.

6.2 PROFILE-LEVEL ANALYSIS

To understand how Stage III improves expected revenue, Table 3 compares the fractions of profiles on which ALGNet, Stage II, and Stage III underperform Item-wise Myerson, together with the fraction on which Stage III underperforms Stage II. As the number of bidders grows, ALGNet falls below

Table 2: Performance comparison of M2L and baseline methods across auction settings. Results for M2L and its intermediate stage are averaged over four independent runs. The standard deviations of M2L revenue, in the order of the reported settings, are 0.005, 0.008, 0.012, 0.018, 0.028, 0.035, 0.036, and 0.033. Regret is reported only for approximately DSIC methods, and dashes indicate results that are not currently reported. The standard deviations of M2L regret, in the order of the reported settings, are 0.0002, 0.0001, 0.0004, 0.0006, 0.0006, 0.0006, 0.0005, and 0.0008.

Class	Method	Metric	2×2	2×5	2×10	3×10	4×10	5×10	6×10	7×10
(1)DD	Item-wise Myerson	Revenue	0.833	2.083	4.164	5.312	6.124	6.718	7.166	7.506
	OD-VVCA	Revenue	0.864	2.255	4.846	**5.817**	**6.484**	**6.986**	7.349	7.662
(2)RD	AMenuNet	Revenue	0.862	2.27	4.78	5.59	6.10	6.45	6.67	6.87
	GemNet	Revenue	0.878	2.310	–	–	–	–	–	–
(3)R	RegretNet	Revenue	0.878	2.189	–	5.541	–	6.778	–	–
		Regret	0.0004	0.0006	–	0.0028	–	0.0048	–	–
	ALGNet	Revenue	0.879	2.183	4.878	5.562	6.323	6.781	7.104	7.403
		Regret	0.0006	0.0007	0.0042	0.0048	0.0054	0.0071	0.0054	0.0055
	After Stage II (Ours)	Revenue	0.882	2.329	4.925	5.730	6.355	6.815	7.219	7.532
		Regret	0.0006	0.0006	0.0042	0.0052	0.0049	0.0043	0.0062	0.0065
	M2L (Ours)	Revenue	**0.885**	**2.345**	**5.044**	5.808	6.455	6.962	**7.362**	**7.691**
		Regret	0.0007	0.0008	0.0049	0.0051	0.0048	0.0051	0.0052	0.0068

Table 3: Profile-level comparison across settings. All Stage-III results are evaluated using the student network alone, without intervention.

Setting	ALGNet below Myerson	Stage II below Myerson	Stage III below Myerson	Stage III below Stage II
2×10	11.40%	8.45%	6.87%	1.12%
3×10	29.08%	14.38%	10.80%	5.33%
4×10	31.18%	25.36%	15.08%	2.75%
5×10	45.78%	40.37%	13.41%	0.09%
6×10	63.74%	46.50%	14.56%	1.15%
7×10	71.32%	35.00%	14.39%	6.89%

Myerson on an increasingly large share of profiles. Stage II consistently reduces these cases, but they remain common in the larger auctions. Stage III further suppresses the below-Myerson fraction across all settings and prevents it from increasing substantially with the number of bidders. Moreover, Stage III rarely reduces revenue relative to Stage II, indicating that its targeted refinement addresses many remaining cases while largely preserving the gains obtained during autonomous exploration.

Runtime details are provided in Appendix C. The Stage-I ablation in Appendix D shows that training Stage II directly from random initialization can be highly unstable in some settings, whereas the Myerson-based initialization provides a reliable starting point. Appendix E controls for the additional training budget by continuing Stage II without Myerson guidance using the Stage-III learning rate, showing that the improvement of Stage III does not arise from longer training alone. The parameter comparison in Appendix F further shows that M2L uses fewer trainable mechanism parameters than ALGNet across all settings.

7 Conclusion

We propose M2L, a three-stage Myerson-guided framework for learning high-revenue multi-item auctions. By combining a Myerson-based structural prior, autonomous multi-item exploration, and selective refinement, M2L retains the flexibility of neural mechanisms while improving their performance relative to Item-wise Myerson. Experiments across different auction scales and valuation distributions show that M2L achieves the highest revenue among direct-parameterization methods while maintaining low regret. It outperforms all baselines in smaller settings and remains competitive with OD-VVCA while substantially outperforming the other methods in larger settings. Future work will investigate how other economic theories and analytically characterized mechanisms can be incorporated into deep-learning-based mechanism design as structural priors or training guidance.

AI USE STATEMENT

In this work, we used generative AI tools for implementing portions of the methods through AI-assisted code generation. We have not used generative AI tools for generating synthetic datasets, developing theoretical models or conceptual frameworks, formulating or proving mathematical claims, designing research methodology or experiments, translating text, cleaning or reformatting datasets, or interpreting results, and qualitative or thematic data analysis and the remaining required-disclosure tasks are not applicable to this work. Additionally, we used generative AI tools for editing the paper to improve readability and generating or editing portions of the software code. We have reviewed all AI-assisted work. All AI-assisted text was reviewed and revised by the authors, and all AI-generated code was manually inspected and tested for correctness by the authors. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Appendix A provides the detailed neural network architecture and design choices of M2L, together with the training and test configurations used in our experiments. The supplementary materials include the complete runnable code and ready-to-use trained PyTorch checkpoint files (`.pt`) for settings with 2–7 bidders and 10 items, enabling direct evaluation without retraining the models.

REFERENCES

Saeed Alaei, Hu Fu, Nima Haghpanah, Jason D. Hartline, and Azarakhsh Malekian. Bayesian optimal auctions via multi- to single-agent reduction. In *Proceedings of the 13th ACM Conference on Electronic Commerce, EC 2012, Valencia, Spain, June 4-8, 2012*, pp. 17, 2012.

Moshe Babaioff, Yannai A Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 869–877, 2017.

Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S Matthew Weinberg. A simple and approximately optimal mechanism for an additive buyer. *Journal of the ACM (JACM)*, 67(4):1–40, 2020.

Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 170–183, 2017.

Michael Curry, Tuomas Sandholm, and John Dickerson. Differentiable economics for randomized affine maximizer auctions. *arXiv preprint arXiv:2202.02872*, 2022.

Zhijian Duan, Jingwu Tang, Yutong Yin, Zhe Feng, Xiang Yan, Manzil Zaheer, and Xiaotie Deng. A context-integrated transformer-based neural network for auction design. In *International Conference on Machine Learning*, pp. 5609–5626. PMLR, 2022.

Zhijian Duan, Haoran Sun, Yurong Chen, and Xiaotie Deng. A scalable neural network for dsic affine maximizer auction design. *Advances in Neural Information Processing Systems*, 36:56169–56185, 2023.

Zhijian Duan, Haoran Sun, Yichong Xia, Siqiang Wang, Zhilin Zhang, Chuan Yu, Jian Xu, and Xiaotie Deng. Automated deterministic auction design with objective decomposition. In *Proceedings of the ACM Web Conference 2026*, pp. 225–236, 2026.

Paul Dütting, Zhe Feng, Harikrishna Narasimhan, David Parkes, and Sai Srivatsa Ravindranath. Optimal auctions through deep learning. In *International Conference on Machine Learning*, pp. 1706–1715, 2019.

Zhe Feng, Harikrishna Narasimhan, and David C Parkes. Deep learning for revenue-optimal auctions with budgets. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, pp. 354–362, 2018.

Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347, 2017.

Sergiu Hart and Noam Nisan. Selling multiple correlated goods: Revenue maximization and menu-size complexity. *Journal of Economic Theory*, 183:991–1029, 2019.

Da Huo, Zhenzhe Zheng, and Fan Wu. Learning optimal auctions with correlated value distributions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 13944–13952, 2025.

Dmitry Ivanov, Iskander Safiulin, Igor Filippov, and Ksenia Balabaeva. Optimal-er auctions through attention. In *NeurIPS*, 2022.

Kevin Kuo, Anthony Ostuni, Elizabeth Horishny, Michael J Curry, Samuel Dooley, Ping-yeh Chiang, Tom Goldstein, and John P Dickerson. Proportionnet: Balancing fairness and revenue for auction design with deep learning. *arXiv preprint arXiv:2010.06398*, 2020.

Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

Gregory Pavlov. Optimal mechanism for selling two goods. *The BE Journal of Theoretical Economics*, 11(1), 2011.

Neehar Peri, Michael Curry, Samuel Dooley, and John Dickerson. Preferencenet: Encoding human preferences in auction design with deep learning. *Advances in Neural Information Processing Systems*, 34:17532–17542, 2021.

Jad Rahme, Samy Jelassi, Joan Bruna, and S Matthew Weinberg. A permutation-equivariant neural network architecture for auction design. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 5664–5672, 2021a.

Jad Rahme, Samy Jelassi, and S. Matthew Weinberg. Auction learning as a two-player game. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*, 2021b.

Weiran Shen, Pingzhong Tang, and Song Zuo. Automated mechanism design via neural networks. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 215–223, 2019.

Tonghan Wang, Yanchen Jiang, and David C Parkes. Gemnet: Menu-based, strategy-proof multi-bidder auctions through deep learning. *arXiv preprint arXiv:2406.07428*, 2024.

Zihe Wang and Pingzhong Tang. Optimal mechanisms with simple menus. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 227–240, 2014.

Andrew Chi-Chih Yao. Dominant-strategy versus bayesian multi-item auctions: Maximum revenue determination and comparison. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pp. 3–20, 2017.

Shuyuan You, Zhiqiang Zhuang, Haiying Wu, Kewen Wang, and Zhe Wang. Improving deep learning powered auction design. In *International Conference on Neural Information Processing*, pp. 556–569. Springer, 2023.

Shuyuan You, Zhiqiang Zhuang, Haiying Wu, Kewen Wang, and Zhe Wang. Bridging the gap between estimated and true regret in deep learning-based auction mechanisms. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems*, pp. 3389–3391, 2026.

A Network Architectures and Hyperparameters

Network Architectures. The allocation, payment, and misreporter networks are fully connected multi-layer perceptrons (MLPs) with Tanh hidden activations. For each item, the allocation network outputs one logit for each bidder and an additional dummy logit representing that the item is not allocated. A softmax over these $n + 1$ outcomes produces bidder allocation probabilities and a no-allocation probability that sum to one. Thus, for every item j, the feasible-allocation constraint $\sum_{i=1}^{n} a_{ij} \leq 1$ is satisfied by construction. The payment and misreporter networks use sigmoid

Table 4: Network architectures used across problem scales. The mechanism column applies to both the allocation and payment networks.

Setting ($n \times m$)	**Mechanism**	**Misreporter**
2×2	5×200	5×200
2×5	5×200	5×200
2×10	7×100	7×100
3×10	7×100	7×100
4×10	7×200	7×200
5×10	8×200	8×200
6×10	8×200	8×200
7×10	8×200	8×200

outputs. Table 4 reports the architectures, denoted by *Depth* $\times$ *Width*, where depth is the number of linear layers.

Training Hyperparameters. We use AdamW with a batch size of 500 and train each stage for 60 epochs. The 2×2 and 2×5 settings use 100,000 samples (200 batches) per epoch; all other settings use 200,000 samples (400 batches) per epoch. Stage I uses mean squared error and learning rate 10^{-3}. In Stage II, both the mechanism and misreporter learning rates are 10^{-3}. In Stage III, the mechanism learning rate is 10^{-4} and the misreporter learning rate is 10^{-3}. In Stages II and III, the misreporter is updated for $R = 100$ inner steps per mechanism update and is re-initialized every two epochs during the first 20 epochs.

Evaluation Hyperparameters. For intermediate evaluation, we optimize 300 randomly initialized misreports for 300 projected Adam steps. Final evaluation uses 10,000 test profiles and follows the item-wise guided refinement approach of You et al. (2026). We first search a coordinate-wise grid of 1,000 candidate values for each bidder–item entry. The resulting initialization pool includes the combined item-wise solution, all single-item deviations, and the best single-item deviation for each bidder. We additionally include one globally random candidate and one Gaussian-perturbed candidate around each of the truthful report and the combined item-wise solution, using a noise standard deviation of 0.2 and clipping reports to $[0, 1]$. Each candidate is refined using 300 projected Adam steps with a learning rate of 10^{-3}. We apply this evaluation configuration consistently to M2L and all reevaluated approximately DSIC baselines.

B Results on Truncated Normal Distribution

We extend our evaluation to a setting where bidder valuations follow a Truncated Normal distribution ($\mu = 0.5, \sigma = 0.5$ within $[0, 1]$). The comparative results are summarized in Table 5.

Table 5: Performance comparison under the truncated normal distribution. Valuations are drawn from a truncated normal distribution with $\mu = 0.5$ and $\sigma = 0.5$ over $[0, 1]$. The standard deviations of After Stage II revenue are 0.014, 0.012, and 0.016 for the 2×5, 3×10, and 5×10 settings, respectively. The standard deviations of M2L revenue are 0.012, 0.014, and 0.018 for the 2×5, 3×10, and 5×10 settings, respectively.

Setting	**Item-wise Myerson**		**ALGNet**		**After Stage II (Ours)**		**M2L (Ours)**	
	Rev.	Reg.	Rev.	Reg.	Rev.	Reg.	Rev.	Reg.
2×5	2.068	0.00	2.201	$< 1 \times 10^{-3}$	2.301	$< 1 \times 10^{-3}$	**2.354**	$< 1 \times 10^{-3}$
3×10	5.225	0.00	5.411	$< 3 \times 10^{-3}$	5.716	$< 3 \times 10^{-3}$	**5.756**	$< 3 \times 10^{-3}$
5×10	6.544	0.00	6.610	$< 5 \times 10^{-3}$	6.632	$< 5 \times 10^{-3}$	**6.810**	$< 5 \times 10^{-3}$

Table 5 shows that the advantage of M2L also holds under the truncated normal distribution. Across all tested settings, both After Stage II and M2L achieve higher revenue than Item-wise Myerson and ALGNet while maintaining low regret. Comparing After Stage II and M2L further shows that Stage III provides an additional and consistent revenue improvement. These results indicate that

the Myerson-guided correction mechanism remains effective when valuations follow a different distribution.

C RUNTIME OF M2L

For completeness, Table 6 reports the runtime of a single complete M2L run in each setting, including all three training stages and the final test-time evaluation. These values are per-run runtimes rather than averages or totals over the four random seeds. As expected, runtime increases with problem scale, mainly due to the higher cost of adversarial misreport optimization in larger settings.

Table 6: Runtime of a single complete M2L run, including training and final evaluation. Runtime is reported in hours.

Setting $(n \times m)$	**Runtime per Run**
2×2	11 h
2×5	12 h
2×10	18 h
3×10	20 h
4×10	28 h
5×10	32 h
6×10	35 h
7×10	39 h

D ABLATION: REMOVING THE STAGE-I INITIALIZATION

Existing neural auction models are commonly trained from random initialization by directly optimizing a joint revenue–regret objective. In challenging multi-item settings, the network must simultaneously learn revenue extraction and approximate DSIC from gradient signals alone. This joint optimization can be highly sensitive to initialization and may converge to qualitatively different mechanisms across random seeds.

We conduct a Stage-I ablation by removing the Myerson-based initialization and training Stage II directly from random initialization using four random seeds (0–3) in the 5×10 setting. This ablation isolates the effect of Stage I on the stability of the subsequent adversarial optimization. Table 7 shows substantial variation across seeds. The runs with seeds 1–3 converge to relatively high-revenue, low-regret mechanisms. In contrast, the run with seed 0 maintains low regret but obtains a revenue of 5.402, below the Item-wise Myerson revenue of 6.718, suggesting convergence to a low-revenue local optimum.

Stage I is designed to reduce this sensitivity by supervising the allocation and payment networks with Item-wise Myerson before the adversarial training in Stage II. Rather than asking a randomly initialized network to learn revenue maximization and approximate DSIC simultaneously, Stage I supplies an economically meaningful mechanism as the starting point. Stage II can subsequently depart from this initialization to explore richer multi-item mechanisms. Thus, Stage I does not restrict the final mechanism to Item-wise Myerson; it provides economic structure that reduces the likelihood of low-revenue local optima during subsequent training. Consistent with this role, we do not observe such a failure in the reported M2L experiments using Stage I. This ablation therefore provides evidence that Stage I improves training stability.

E ABLATION: CONTINUED STAGE-II TRAINING WITHOUT MYERSON GUIDANCE

To determine whether the gains of Stage III arise from Myerson guidance or simply from additional optimization, we remove the Myerson-guided intervention from Stage III. Starting from the same Stage-II checkpoint, we continue optimizing the original Stage-II revenue–regret objective for the same number of updates as Stage III, while using the Stage-III mechanism learning rate of 10^{-4}. All

Table 7: Stage-I ablation: revenue and regret when Stage II is trained directly from random initialization using seeds 0–3 in the 5×10 setting. Boldface denotes the run with revenue below Item-wise Myerson.

Random Seed	Revenue	Regret
0	**5.402**	2.01×10^{-3}
1	6.776	3.37×10^{-3}
2	6.785	3.01×10^{-3}
3	6.757	3.01×10^{-3}

other training settings are kept unchanged. This ablation matches the additional training budget and learning-rate schedule of Stage III while removing its Myerson-guided training signal.

Table 8: Revenue comparison between the Stage-II checkpoint, continued Stage-II training without Myerson guidance, and the full M2L framework.

Setting $(n \times m)$	**Stage II**	**Continued Stage II**	**M2L**
2×2	0.882	0.881	0.885
2×5	2.329	2.326	2.345
2×10	4.925	4.926	5.044
3×10	5.730	5.716	5.808
4×10	6.355	6.356	6.455
5×10	6.815	6.818	6.962
6×10	7.219	7.224	7.362
7×10	7.532	7.531	7.691

As shown in Table 8, continued Stage-II training provides no meaningful revenue improvement over the original Stage-II checkpoint. Under the tested training schedule, additional optimization of the same objective, even with the smaller Stage-III learning rate, yields only minor revenue changes. In contrast, the full M2L framework improves revenue by introducing the Myerson-guided training signal. The benefit of Stage III therefore comes from its structural guidance rather than from additional training time or a different learning-rate schedule.

F Parameter Count Comparison with ALGNet

We compare the number of trainable mechanism parameters between ALGNet and M2L in Table 9. For ALGNet, the auctioneer consists of two allocation-related MLPs and one payment MLP. In contrast, the M2L mechanism consists of one allocation network and one payment network.

Table 9: Number of trainable mechanism parameters in ALGNet and M2L.

Setting $(n \times m)$	**ALGNet Mechanism**	**M2L Mechanism**
2×2	365,805	244,004
2×5	370,611	247,610
2×10	159,921	107,220
3×10	162,921	109,220
4×10	631,821	422,420
5×10	758,421	506,820
6×10	764,421	510,820
7×10	1,153,821	514,820

Across all settings, M2L uses fewer mechanism parameters than ALGNet.