A Permutation-Equivariant Neural Network Architecture For Auction Design

Jad Rahme[1], **Samy Jelassi** [1,2], **Joan Bruna** [2,3], **S. Matthew Weinberg** [1]
[1]Princeton University, USA
[2]Courant Institute of Mathematical Sciences, New York University, USA
[3]Center for Data Science, New York University, USA
jrahme@princeton.edu

Abstract

Designing an incentive compatible auction that maximizes expected revenue is a central problem in Auction Design. Theoretical approaches to the problem have hit some limits in the past decades and analytical solutions are known for only a few simple settings. Building on the success of deep learning, a new approach was recently proposed by Duetting et al. (2019) in which the auction is modeled by a feed-forward neural network and the design problem as a learning problem. However, the architectures used in that work are general purpose and do not take advantage of any structure the solution might possess. For example, *symmetric* auctions are known to be optimal in many settings of interest, and near-optimal quite generally (Daskalakis and Weinberg 2012; Kothari et al. 2019a), yet previous architectures do not recover this structure (even in settings where it is known to exist). In this work, we construct a neural architecture that is capable of perfectly recovering the optimal *symmetric* mechanism. We further demonstrate that permutation-equivariant architectures are not only capable of recovering previous results, they also have better generalization properties.

1 Introduction

Designing truthful auctions is one of the core problems that arise in economics. Concrete examples of auctions include sales of treasury bills, art sales by Christie's or Google Ads. Following seminal work of Vickrey (Vickrey 1961) and Myerson (Myerson 1981), auctions are typically studied in the *independent private valuations* model: each bidder has a valuation function over items, and their payoff depends only on the items they receive. Moreover, the auctioneer knows aggregate information about the population that each bidder comes from, modeled as a distribution over valuation functions, but does not know precisely each bidder's valuation. Auction design is challenging since the valuations are private and bidders need to be encouraged to report their valuations truthfully. The auctioneer aims at designing an incentive compatible auction that maximizes revenue.

While auction design has existed as a subfield of economic theory for several decades, complete characterizations of the optimal auction only exist for a few settings. Myerson resolved the optimal auction design problem when there is a single item for sale (Myerson 1981). However, the problem is not completely understood even in the extremely simple setting with just a single bidder and two items. While there have been some partial characterizations (Manelli and Vincent 2006, 2010; Pavlov 2011; Wang and Tang 2014; Daskalakis, Deckelbaum, and Tzamos 2017), and algorithmic solutions with provable guarantees (Alaei 2011; Alaei et al. 2012, 2013; Cai, Daskalakis, and Weinberg 2012a,b), neither the analytic nor algorithmic approach currently appears tractable for seemingly small instances.

Another line of work to confront this theoretical hurdle consists in building automated methods to find the optimal auction. Early works (Conitzer and Sandholm 2002, 2004) framed the problem as a linear program. However, this approach suffers from severe scalablility issues as the number of constraints and variables is exponential in the number of bidders and items (Guo and Conitzer 2010). Later, Sandholm and Likhodedov (2015) designed algorithms to find the optimal auction. While scalable, they are however limited to specific classes of auctions known to be incentive compatible. A more recent research direction consists in building deep learning architectures that design auctions from samples of bidder valuations. Duetting et al. (2019) proposed RegretNet, a feed-forward architecture to find near-optimal results in several known multi-item settings and obtain new mechanisms in unknown cases. This architecture however is not data efficient and can require a large number of valuation samples to learn an optimal auction in some cases. This inefficiency is not specific to RegretNet but is characteristic of neural network architectures that do not incorporate any inductive bias.

In this paper, we build a deep learning architecture for *multi-bidder symmetric auctions*. These are auctions which are invariant to relabeling the items or bidders. More specifically, such auctions are *anonymous* (in that they can be executed without any information about the bidders, or labeling them) and *item-symmetric* (in that it only matters what bids are made for an item, and not its a priori label).

It is now well-known that when bidders come from the same population that the optimal auction itself is anonymous. Similarly, if items are *a priori indistinguishable* (e.g. different colors of the same car — individuals certainly value a red vs. blue car differently, but there is nothing objectively more/less valuable about a red vs. blue car), the optimal auction is itself item-symmetric. In such settings, our approach will approach the true optimum in a way which retains this structure (see Contributions below). Even without these conditions, the optimal auction is often symmetric anyway: for example, “bundling together” (the auction which allows bidders to pay a fixed price for all items, or receive nothing) is item-symmetric, and is often optimal even when the items are a priori distinguishable.

Beyond their frequent optimality, such auctions are desirable objects of study *even when they are suboptimal*. For example, seminal work of Hartline and Roughgarden which pioneered the study of “simple vs. optimal auctions” analyzes the approximation guarantees achievable by anonymous auctions (Hartline and Roughgarden 2009), and exciting recent work continues to improve these guarantees (Alaei et al. 2015; Jin et al. 2019b,a). Similarly, Daskalakis and Weinberg (2012) develop algorithms for item-symmetric instances, and exciting recent work show how to leverage item-symmetric to achieve near-optimal auctions in completely general settings (Kothari et al. 2019b). To summarize: symmetric auctions are known to be optimal in many settings of interest (even those which are not themselves symmetric). Even in settings where they are not optimal, they are known to yield near-optimal auctions. And even when they are only approximately optimal, seminal work has identified them as important objects of study owing to their simplicity. In modern discussion of auctions, they are also desirable due to fairness considerations.

While applying existing feed-forward architectures as RegretNet to symmetric auctions is possible, we show in Section 3 that RegretNet struggles to find symmetric auctions, *even when the optimum is symmetric*. To be clear, the architecture's performance is indeed quite close to optimal, but the resulting auction is not “close to symmetric”. This paper proposes an architecture that outputs a symmetric auction symmetry by design.

Contributions

This paper identifies three drawbacks from using the RegretNet architecture when learning with symmetric auctions. First, RegretNet is incapable of finding symmetric auctions when the optimal mechanism is known to be symmetric. Second, RegretNet is sample inefficient, which is not surprising since the architecture does not incorporate any inductive bias. Third, RegretNet is incapable of generalizing to settings with a different number of bidders of objects. In fact, by construction, the solution found by RegretNet can only be evaluated on settings with exactly the same number of bidders and objects of the setting it was trained on.

We address these limitations by proposing a new architecture EquivariantNet, that outputs symmetric auctions. EquivariantNet is an adaption of the deep sets architecture (Hartford et al. 2018) to symmetric auctions. This architecture is parameter-efficient and is able to recover some of the optimal results in the symmetric auctions literature. Our approach outlines three important benefits:

– *Symmetry*: our architecture outputs a symmetric auction by design. It is immune to permutation-sensitivity as defined in Section 3 which is related to fairness.

– *Sample generalization*: Because we use domain knowledge, our architecture converges to the optimum with fewer valuation samples.

– *Out-of-setting generalization*: Our architecture does not require hard-coding the number of bidders or items during training — training our architecture on instances with n bidders and m items produces a well-defined auction even for instances with n' bidders and m' items. Somewhat surprisingly, we show in 4 some examples where our architecture trained on 1 bidder with 5 items generalizes well even to 1 bidder and m items, for any $m \in \{2, 10\}$.

We highlight that the novelty of this paper is not to show that a new architecture is a viable alternative to RegretNet. Instead we are solving three fundamental limitations we identified for the RegretNet architecture. These three problems are not easy to solve in principle, it is surprising that a change of architecture solves all of them in the context of symmetric auctions. We would also like to emphasize that both RegretNet and EquivariantNet are capable of learning auction with near optimal revenue and negligible regret. It is not possible to significantly outperform RegretNet on these aspects. The way we improve over RegretNet is by having better sample efficiency, out-of-setting generalization and by ensuring that our solutions are exactly equivariant.

The paper decomposes as follows. Section 2 introduces the standard notions of auction design. Section 3 presents our permutation-equivariant architecture to encode symmetric auctions. Finally, Section 4 presents numerical evidence for the effectiveness of our approach.

Related Work

Auction design and machine learning. Machine learning and computational learning theory have been used in several ways to design auctions from samples of bidder valuations. Some works have focused sample complexity results for designing optimal revenue-maximizing auctions. This has been established in single-parameter settings (Dhangwatnotai, Roughgarden, and Yan 2015; Cole and Roughgarden 2014; Morgenstern and Roughgarden 2015; Medina and Mohri 2014; Huang, Mansour, and Roughgarden 2018; Hartline and Taggart 2019; Gonczarowski and Nisan 2017; Guo, Huang, and Zhang 2019), multi-item auctions (Dughmi, Han, and Nisan 2014; Gonczarowski and Weinberg 2018), combinatorial auctions (Balcan, Sandholm, and Vitercik 2016; Morgenstern and Roughgarden 2016) and allocation mechanisms (Narasimhan and Parkes 2016). Machine learning has also been used to optimize different aspects of mechanisms (Lahaie 2011; Dütting et al. 2015). All these aforementioned differ from ours as we resort to deep learning for finding optimal auctions.

Auction design and deep learning. While Duetting et al. (2019) is the first paper to design auctions through deep learning, several other paper followed-up this work. (Feng, Narasimhan, and Parkes 2018) extended it to budget constrained bidders, (Golowich, Narasimhan, and Parkes 2018) to the facility location problem. Tacchetti et al. (2019) built architectures based on the Vickrey-Clarke-Groves auctions.

Rahme, Jelassi, and Weinberg (2021) proposed a new formulation of the auction learning problem as a two player game. Recently, Shen, Tang, and Zuo (2019) and Duetting et al. (2019) proposed architectures that *exactly* satisfy incentive compatibility but are specific to *single-bidder* settings. In this paper, we aim at *multi-bidder* settings and build permutation-equivariant networks that return nearly incentive compatibility symmetric auctions. Lastly, we would like to mention that a concurrent work (Duetting et al. 2019) reported in a recent version of their preprint achieving faster training times by imposing similar symmetries in their architecture. However, they did not provide any details on their approach nor exhibited their numerical performance.

2 Symmetries and Learning Problem in Auction Design

We review the framework of auction design and the problem of finding truthful mechanisms. We then present symmetric auctions and similarly to Duetting et al. (2019), frame auction design as a learning problem.

Auction Design and Symmetries

Auction design. We consider the setting of additive auctions with n bidders with $N = \{1, \dots, n\}$ and m items with $M = \{1, \dots, m\}$. Each bidder i is has value v_{ij} for item j, and values the set S of items at $\sum_{j \in S} v_{ij}$. Such valuations are called *additive*, and are perhaps the most well-studied valuations in multi-item auction design (Hart and Nisan 2012, 2013; Li and Yao 2013; Babaioff et al. 2014; Hart and Reny 2015; Daskalakis, Deckelbaum, and Tzamos 2017; Beyhaghi and Weinberg 2019).

The designer does not know the full valuation profile $V = (v_{ij})_{i \in N, j \in M}$, only the distribution from which they are drawn. Specifically, the valuation vector of bidder i for each of the m items $\vec{v}_i = (v_{i1}, \dots, v_{im})$ is drawn from a distribution D_i over $\mathbb{R}^m$ (V is then drawn from $D := \times_i D_i$). The designer asks the bidders to report their valuations (potentially untruthfully), then decides on an allocation of items to the bidders and charges a payment to them.

Definition 2.1 *An auction is a pair (g, p) consisting of a randomized allocation rule $g = (g_1, \dots, g_n)$ where $g_i \colon \mathbb{R}^{n \times m} \to [0,1]^m$ such that for all V, and all j, $\sum_i (g_i(V))_j \leqslant 1$ and payment rules $p = (p_1, \dots, p_n)$ where $p_i \colon \mathbb{R}^{n \times m} \to \mathbb{R}_{\geqslant 0}$.*

Given reported bids $B = (b_{ij})_{i \in N, j \in M}$, the auction computes an allocation probability $g(B)$ and payments $p(B)$. $[g_i(B)]_j$ is the probability that bidder i received object j and $p_i(B)$ is the price bidder i has to pay to the mechanism. In what follows, $\mathcal{M}$ denotes the class of all possible auctions.

Definition 2.2 *The utility of bidder i is defined by $u_i(\vec{v}_i, B) = \sum_{j=1}^m [g_i(B)]_j v_{ij} - p_i(B)$.*

Bidders seek to maximize their utility and may report bids that are different from their valuations. Let V_{-i} be the valuation profile without element $\vec{v}_i$, similarly for B_{-i} and $D_{-i} = \times_{j \neq i} D_j$. We aim at auctions that invite bidders to bid their true valuations through the notion of incentive compatibility.

Definition 2.3 *An auction (g, p) is dominant strategy incentive compatible (DSIC) if each bidder's utility is maximized by reporting truthfully no matter what the other bidders report. For every bidder i, valuation $\vec{v}_i \in D_i$, bid $\vec{b}_i{}' \in D_i$ and bids $B_{-i} \in D_{-i}$, $u_i(\vec{v}_i, (\vec{v}_i, B_{-i})) \geqslant u_i(\vec{v}_i, (\vec{b}_i{}', B_{-i}))$.*

Additionally, we aim at auctions where each bidder receives a non-negative utility.

Definition 2.4 *An auction is individually rational (IR) if for all $i \in N$, $\vec{v}_i \in D_i$ and $B_{-i} \in D_{-i}$,*

$$u_i(\vec{v}_i, (\vec{v}_i, B_{-i})) \geqslant 0. \tag{IR}$$

In a DSIC auction, the bidders have the incentive to truthfully report their valuations and therefore, the revenue on valuation profile V is defined as $\sum_{i=1}^n p_i(V)$. Optimal auction design aims at finding a DSIC auction that maximizes the expected revenue $rev := \mathbb{E}_{V \sim D}[\sum_{i=1}^n p_i(V)]$.

Linear program. We frame the problem of optimal auction design as an optimization problem where we seek an auction that minimizes the negated expected revenue among all IR and DSIC auctions. Since there is no known characterization of DSIC mechanisms in the multi-bidder setting, we resort to the relaxed notion of *ex-post regret*. It measures the extent to which an auction violates DSIC, for each bidder.

Definition 2.5 *The ex-post regret for a bidder i is the maximum increase in his utility when considering all his possible bids and fixing the bids of others. For a valuation profile V, the ex-post regret for a bidder i is $rgt_i(V) = \max_{\vec{v}_i{}' \in \mathbb{R}^m} u_i(\vec{v}_i; (\vec{v}_i{}', V_{-i})) - u_i(\vec{v}_i; (\vec{v}_i, V_{-i}))$. In particular, DSIC is equivalent to*

$$rgt_i(V) = 0, \ \forall i \in N. \tag{IC}$$

Therefore, by setting (IC) and (IR) as constraints, finding an optimal auction is equivalent to the following linear program

$$\min_{(g,p) \in \mathcal{M}} - \mathbb{E}_{V \sim D}\left[\sum_{i=1}^n p_i(V)\right] \quad \text{such that}$$

$$\begin{cases} rgt_i(V) = 0 & \forall i \in N, \ \forall V \in D, \\ u_i(\vec{v}_i, (\vec{v}_i, B_{-i})) \geqslant 0, & \forall i \in N, \ \vec{v}_i \in D_i, B_{-i} \in D_{-i} \end{cases} \tag{LP}$$

Symmetric auctions. (LP) is intractable due to the exponential number of constraints. However, in the setting of *symmetric* auctions, it is possible to reduce the search space of the problem as shown in 2.1. We first define the notions of bidder- and item-symmetries.

Definition 2.6 *The valuation distribution D is bidder-symmetric if for any permutation of the bidders $\varphi_b \colon N \to N$, the permuted distribution $D_{\varphi_b} := D_{\varphi_b(1)} \times \cdots \times D_{\varphi_b(n)}$ satifies: $D_{\varphi_b} = D$.*

Bidder-symmetry intuitively means that the bidders are a priori indistinguishable (although individual bidders will be different). This holds for instance in auctions where the identity of the bidders is anonymous, or if $D_i = D_j$ for all i, j (bidders are i.i.d.).

Definition 2.7 *Bidder i's valuation distribution D_i is item-symmetric if for any items $x_1, \dots, x_m$ and any permutation $\varphi_o \colon M \to M$, $D_i(x_{\varphi_o(1)}, \dots, x_{\varphi_o(m)}) = D_i(x_1, \dots, x_m)$.*

Intuitively, item-symmetry means that the items are also indistinguishable but not identical. It holds when the distributions over the items are i.i.d. but this is not a necessary condition. Indeed, the distribution $\{(a, b, c) \in \mathcal{U}(0, 1)^{\otimes 3} : a + b + c = 1\}$ is not i.i.d. but is item-symmetric.

Definition 2.8 *An auction is symmetric if its valuation distributions are bidder- and item-symmetric.*

We now define the notion of permutation-equivariance that is important in symmetric auctions.

Definition 2.9 *The functions g and p are permutation-equivariant if for any two permutation matrices $\Pi_n \in \{0,1\}^{n\times n}$ and $\Pi_m \in \{0,1\}^{m\times m}$, and any valuation matrix V, we have $g(\Pi_n\, V\, \Pi_m) = \Pi_n\, g(V)\, \Pi_m$ and $p(\Pi_n\, V\, \Pi_m) = \Pi_n\, p(V)$.*

Theorem 2.1 *When the auction is symmetric, there exists an optimal solution to* (LP) *that is permutation-equivariant.*

Theorem 2.1 is originally proved in Daskalakis and Weinberg (2012) and its proof is reminded in Appendix B for completeness. It encourages to reduce the search space in LP by only optimizing over permutation-equivariant allocations and payments. We implement this idea in Section 3 where we build equivariant neural network architectures. Before, we frame auction design as a learning problem.

Auction Design as a Learning Problem

Similarly to Duetting et al. (2019), we formulate auction design as a learning problem. We learn a parametric set of auctions (g^w, p^w) where $w \in \mathbb{R}^d$ parameters and $d \in \mathbb{N}$. Directly solving (LP) is challenging in practice. Indeed, the auctioneer must have access to the bidder valuations which are unavailable to her. Since she has access to the valuation distribution, we relax (LP) and replace the IC constraint for all $V \in D$ by the expected constraint $\mathbb{E}_{V\sim D}[rgt_i(V)] = 0$ for all $i \in N$.. In practice, the expectation terms are computed by sampling L bidder valuation profiles drawn i.i.d. from D. The empirical ex-post regret for bidder i is

$$\widehat{rgt}_i(w) = \frac{1}{L}\sum_{\ell=1}^{L} \max_{\vec{v}_i{}' \in \mathbb{R}^m} u_i^w(\vec{v}_i^{(\ell)}; (\vec{v}_i{}', V_{-i}^{(\ell)})) - u_i(\vec{v}_i^{(\ell)}; (\vec{v}_i^{(\ell)}, V_{-i}^{(\ell)})), \tag{$\hat{R}$}$$

where $u_i^w(\vec{v}_i, B) := \sum_{j=1}^m [g_i^w(B)]_j v_{ij} - p_i^w(B)$ is the utility of bidder i under the parametric set of auctions (g^w, p^w). Therefore, the learning formulation of (LP) is

$$\min_{w\in\mathbb{R}^d} -\frac{1}{L}\sum_{\ell=1}^{L}\sum_{i=1}^{n} p_i^w(V^{(\ell)}) \quad \text{s.t.} \quad \widehat{rgt}_i(w) = 0, \tag{$\widehat{\mathrm{LP}}$}$$

Duetting et al. (2019) justified the validity of this reduction from (LP) to $(\widehat{\mathrm{LP}})$ by showing that the gap between the expected regret and the empirical regret is small as the number of samples increases. Additionally to being DSIC, the auction must satisfy IR. The learning problem $(\widehat{\mathrm{LP}})$ does not ensure this but we will show how to include this requirement in the architecture in Section 3.

3 Permutation-Equivariant Neural Network Architecture

We first show that fully connected architectures as RegretNet (Duetting et al. 2019) may struggle to find a symmetric solution in auctions where the optimal solution is known to be symmetric. We then describe our neural network architecture, EquivariantNet that learns symmetric auctions. EquivariantNet is build using exchangeable matrix layers (Hartford et al. 2018).

Feed-Forward Nets and Permutation-Equivariance

In the following experiments we use the RegretNet architecture with the exact same training procedure and parameters as found in Duetting et al. (2019).

Permutation-sensitivity. Given L bidders valuation samples $\{B^{(1)}, \dots, B^{(L)}\} \in \mathbb{R}^{n\times m}$, we generate for each bid matrix $B^{(\ell)}$ all its possible permutations $B^{(\ell)}_{\Pi_n,\Pi_m} := \Pi_n B^{(\ell)} \Pi_m$, where $\Pi_n \in \{0,1\}^{n\times n}$ and $\Pi_m \in \{0,1\}^{m\times m}$ are permutation matrices. We then compute the revenue for each one of these bid matrices and obtain a revenue matrix $R \in \mathbb{R}^{n!m!\times L}$. Finally, we compute $h_R \in \mathbb{R}^L$ where $[h_R]_j = \max_{i\in[n!m!]} R_{ij} - \min_{i\in[n!m!]} R_{ij}$. The distribution given by the entries of h_R is a measure of how close the auction is to permutation-equivariance. A symmetric mechanism satisfies $h_R = (0, \dots, 0)^\top$. Our numerical investigation considers the following auction settings:

(I) One bidder, two items. The item values are drawn from $\mathcal{U}[0, 1]$. Optimal revenue: 0.55 (Manelli and Vincent 2006).

(II) Four bidders, five items. Item values are drawn independently from $\mathcal{U}[0, 1]$.

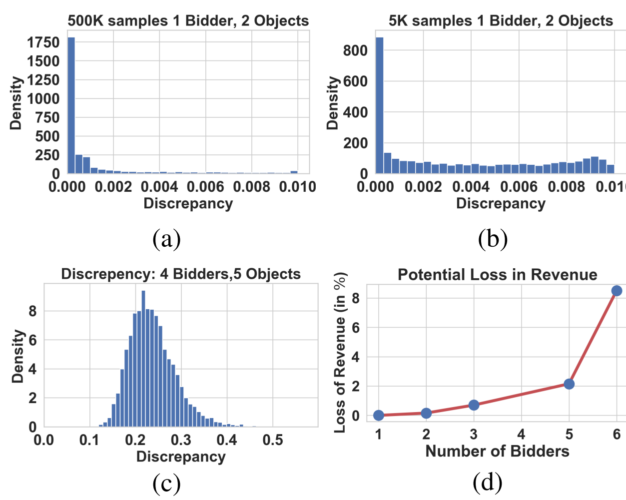

Figure 1: (a)-(b): Distribution h_R when varying the number of training samples (a) 500 000 (b) 5000 samples. (c): Histogram of the distribution h_R for setting (II). (d): Maximum revenue loss as a function of n for setting (III$_n$).

Figure 1 (a)-(b) presents the distribution of h_R of the optimal auction learned for setting (I) when varying the number of samples L. When L is large, the distribution is almost concentrated at zero and therefore the network is almost able to recover the permutation-equivariant solution. When L is small, h_R is less concentrated around zero and therefore, the solution obtained is non permutation-equivariant.

As the problem's dimensions increase, this lack of permutation-invariance becomes more dramatic. Figure 1 (c) shows h_R for the optimal auction mechanism learned for setting (II) when trained with $5 \cdot 10^5$ samples. Contrary to (I), almost no entry of h_R is located around zero, they are concentrated around between 0.1 and 0.4 i.e. between 3.8% and 15% of the estimated optimal revenue.

Exploitability. To highlight how important equivariant solutions are, we analyze the worst-revenue loss that the auctioneer can incur when the bidders act adversarially. Indeed, since different permutations can result in different revenues for the auction, cooperative bidders could pick among the $n!$ possible permutations of their labels the one that minimized the revenue of the mechanism and present themselves in that order. Instead of getting a revenue of $R_{opt} = \mathbb{E}_{V\sim D}\left[\sum_{i=1}^n p_i(V)\right]$, the auctioneer would get a revenue of $R_{adv} = \mathbb{E}_{V\sim D}\left[\min_{\Pi_n}\{\sum_{i=1}^n p_i(\Pi_n V)\}\right]$. The percentage of revenue loss is given by $l = 100 \times \frac{R_{opt}-R_{adv}}{R_{opt}}$. We compute l in in the following family of settings:

(III_n) n additive bidders and ten item where the item values are drawn from $\mathcal{U}[0,1]$.

In Figure 1(d) we plot $l(n)$, the loss in revenue as a function of n. As n increases, $l(n)$ becomes more substantial getting over the 8% with only $n = 6$ bidders.

While it is unlikely that all the bidders will collide and exploit the bidding mechanism in real life, these investigations of permutation sensitivity and exploitability give us a sense of how far the solutions found by RegretNet are from being bidder-symmetric. The underlying real problem with non bidder-symmetric solution has to do with fairness. RegretNet finds mechanisms that do not treat all bidders equally. Their row number in the bid matrix matters, two bidders with the same bids will not get the same treatment. If the mechanism is equivariant however, all bidders will be treated equally by design, there are no biases or special treatments. Aiming for symmetric auctions is important and to this end, we design a permutation-equivariant architecture.

Architecture for Symmetric Auctions (EquivariantNet)

Our input is a bid matrix $B = (b_{i,j}) \in \mathbb{R}^{n\times m}$ drawn from a bidder-symmetric and item-symmetric distribution. We aim at learning a randomized allocation neural network $g^w\colon \mathbb{R}^{n\times m} \to [0,1]^{n\times m}$ and a payment network $p^w\colon \mathbb{R}^{n\times m} \to \mathbb{R}^n_{\geqslant 0}$. The symmetries of the distribution from which B is drawn and Theorem 2.1 motivates us to model g^w and p^w as permutation-equivariant functions. To this end, we use *exchangeable matrix layers* (Hartford et al. 2018) and their definition and properties are reminded in Appendix A. We now describe the three modules of the allocation and payment networks Figure 2.

The first network outputs a vector $q^w(B) \in [0,1]^m$ such that entry $q^w_j(B)$ is the probability that item j is allocated to any of the n bidders. The architecture consists of three modules. The first one is a deep permutation-equivariant network with tanh activation functions. The output of that module is a matrix $Q \in \mathbb{R}^{n\times m}$. The second module transforms Q into a vector $\mathbb{R}^m$ by taking the average over the rows of Q. We finally apply the sigmoid function to the result to ensure that $q^w(B) \in [0,1]^m$. This architecture ensures that $q^w(B)$ is invariant with respect to bidder permutations and equivariant with respect to items permutations.

The second network outputs a matrix $h(B) \in [0,1]^{n\times m}$ where h^w_{ij} is the probability that item j is allocated to bidder i conditioned on item j being allocated. The architecture consists of a deep permutation-equivariant network with tanh activation functions followed by softmax activation function so that $\sum_{i=1}^n h^w_{ij}(B) = 1$. This architecture ensures that q^w equivariant with respect to object and bidder permutations.

By combining the outputs of q^w and h^w, we compute the allocation function $g^w\colon \mathbb{R}^{n\times m} \to [0,1]^{n\times m}$ where $g_{ij}(B)$ is the probability that the allocated item j is given to bidder i. Indeed, using conditional probabilities, we have $g^w_{ij}(B) = q^w_j(B)h^w_{ij}(B)$. g^w is a permutation-equivariant function.

The third network outputs a vector $p(B) \in \mathbb{R}^n_{\geqslant 0}$ where $\tilde{p}^w_i$ is the fraction of bidder's i utility that she has to pay to the mechanism. Given the allocation function g^w, bidder i has to pay an amount $p_i = \tilde{p}_i(B)\sum_{j=1}^m g^w_{ij}(B)B_{ij}$. Individual rationality is ensured by having $\tilde{p}_i \in [0,1]$. The architecture of $\tilde{p}^w$ is almost similar to the one of q^w. Instead of averaging over the rows of the matrix output by the permutation-equivariant architecture, we average over the columns.

Optimization and Training

The optimization and training procedure of EquivariantNet is similar to Duetting et al. (2019). For this reason, we briefly mention the outline of this procedure and remind the details in Appendix C. We apply the augmented Lagrangian method to ($\hat{R}$). The Lagrangian with a quadratic penalty is:

$$\mathcal{L}_\rho(w;\lambda) = -\frac{1}{L}\sum_{\ell=1}^{L}\sum_{i\in N} p^w_i(V^{(\ell)}) + \sum_{i\in N}\lambda_i \widehat{rgt}_i(w) + \frac{\rho}{2}\left(\sum_{i\in N}\widehat{rgt}_i(w)\right)^2$$

where $\lambda \in \mathbb{R}^n$ is a vector of Lagrange multipliers and $\rho > 0$ is a fixed parameter controlling the weight of the quadratic penalty. The solver alternates between the updates on model parameters and Lagrange multipliers: $w^{new} \in \text{argmax}_w \mathcal{L}_\rho(w^{old}, \lambda^{old})$ and $\lambda^{new}_i = \lambda^{old}_i + \rho\cdot\widehat{rgt}_i(w^{new})$.

4 Experimental Results

We first show the effectiveness of EquivariantNet in symmetric and asymmetric auctions. We then highlight its sample-efficiency and ability to extrapolate to other settings.

Evaluation. In addition to the revenue of the learned auction, we also evaluate the empirical average regret over bid-

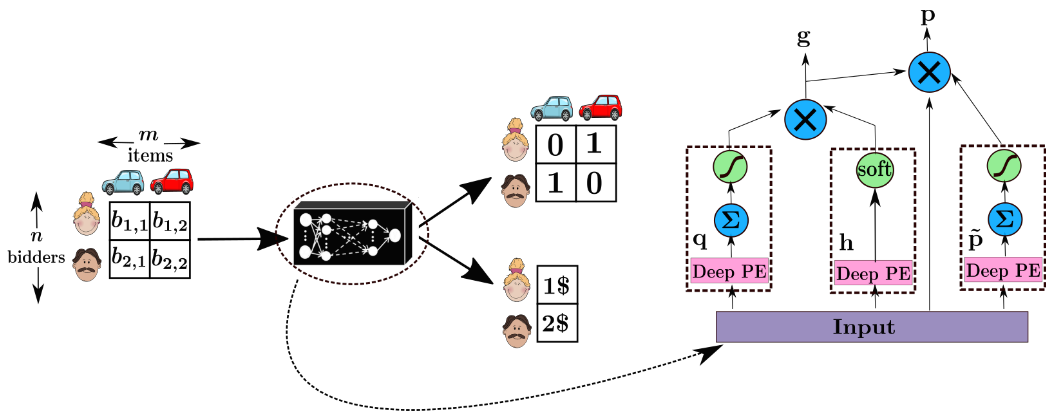

Figure 2: *Left:* Auction design setting. *Right:* EquivariantNet: Deep permutation-equivariant architecture for auction design. Deep PE denotes the deep permutation-equivariant architecture described in Section 3. The symbol soft stands for softmax and the curve for sigmoid. The network outputs an allocation g and a payment p.

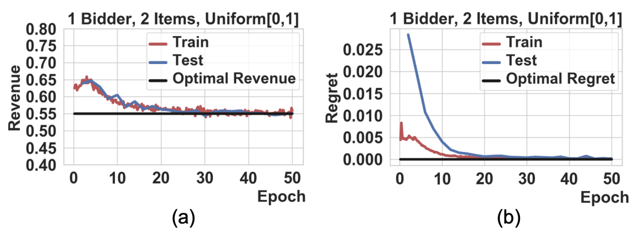

Figure 3: Train/test revenue (a) and regret (b) as a function of epochs for setting (I) for EquivariantNet. The revenue converges to the theoretical optimum (0.55) and the regret converges to 0. EquivariantNet recovers the optimal mechanism.

Dist.	rev	rgt	OPT
(I)	0.551	0.00013	0.550
(IV)	0.173	0.00003	0.1706
(V)	0.873	0.001	0.860

Table 1: Test revenue and regret found by EquivariantNet for settings (I), (IV) and (V). For seeting (V) OPT is the optimal revenue from VVCA and AMA_{bsym} families of auctions (Sandholm and Likhodedov 2015). For settings (I) and (IV), OPT is the theoretical optimal revenue.

ders $\widehat{rgt} = \frac{1}{n}\sum_{i=1}^{n}\widehat{rgt}_i$. We evaluate these terms by running gradient ascent on v_i' with a step-size of 0.001 for $\{300, 500\}$ iterations (we test $\{100, 300\}$ different random initial v_i' and report the one achieves the largest regret).

Known optimal solution. We consider single bidder multi-item auctions where the optimal auction is known to be symmetric. This includes independent private value auctions such as (I) but, surprisingly, some item-asymmetric auction such as (IV) have also an optimal symmetric solution:
(IV) One bidder and two items where the item values are independently drawn according to the probability densities $f_1(x) = 5/(1+x)^6$ and $f_2(y) = 6/(1+y)^7$. Optimal solution in Daskalakis, Deckelbaum, and Tzamos (2017).

The first two lines in Table 1 report the revenue and regret of the learned mechanism for settings (I) and (IV). The revenue is very close to the optimal one, and the regret is negligible. The learned auction may achieve a revenue slightly above the optimal incentive compatible auction. This is possible because although small, the regret is non-zero.

Figure 3(a)-(b) presents a plot of revenue and regret as a function of training epochs for the setting (I). We can see that the revenue converges to the theoretical optimum while the regret converges to zero on both the training and testing set. EquivariantNet recovered the optimal mechanism.

Unknown optimal solution. Our architecture is also able to recover a permutation-equivariant solution in settings for which the optimum is not known analytically such as:

(V) Two additive bidders and two items where bidders draw their value for each item from $\mathcal{U}[0, 1]$.

We compare our solution to the optimal auctions from the VVCA and AMA_{bsym} families of incentive compatible auctions (Sandholm and Likhodedov 2015). The last line of Table 1 summarizes our results. We can see that EquivariantNet finds auctions that are competitive with previous methods.

Non-symmetric optimal solution. When the auction is non symmetric, modeling the solution with permutation-equivariant solutions can in principle lead arbitrarily bad results. Nonetheless, in many cases, equivariant solutions can approximate non equivariant optimal solution very well. Here we show that EquivariantNet is capable of returning satisfactory results in asymmetric auctions. (VI) is a setting where there may not be permutation-equivariant solutions.

(VI) Two bidders, two items. The item values are independently drawn according to the probability densities $f_1(x) = \lambda_1^{-1}e^{-\lambda_1 x}$ and $f_2(y) = \lambda_2^{-1}e^{-\lambda_2 y}$, where $\lambda_1, \lambda_2 > 0$.

Table 2 shows the revenue and regret of the final auctions learned for setting (VI). When $\lambda_1 = \lambda_2$, the auction is symmetric, the revenue of the learned auction is very close to the optimal revenue, with negligible regret. As we increase the gap between λ_1 and λ_2, the asymmetry becomes dom-

λ_2	EquivariantNet rev	rgt	RegretNet rev_F	rgt_F
0.01	0.37	0.0006	0.39	0.0003
0.1	0.41	0.0004	0.41	0.0007
1	0.86	0.0005	0.84	0.0012
10	3.98	0.0081	3.96	0.0056

Table 2: Test revenue/regret for setting (VI) when varying λ_2 ($\lambda_1 = 1$). rev_F and rgt_F are computed with RegretNet.

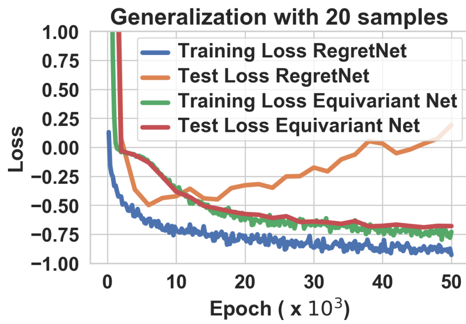

Figure 4: Train and test losses (the Lagrangian) for setting (V) with 20 training samples. RegretNet and EquivariantNet both achieve small losses on the training set, only EquivariantNet is generalizes to the testing set.

inant and the optimal auction does not satisfy permutation-equivariance. We remark that our architecture does output a solution with near-optimal revenue and small regret.

Sample-efficiency. Our permutation-equivariant architecture exhibits solid generalization properties when compared to the feed-forward architecture RegretNet. When enough data is available at training, both architectures generalize well to unseen data and the gap between the training and test losses goes to zero. However, when fewer training samples are available, our equivariant architecture generalizes while RegretNet struggles to. This may be explained by the inductive bias in our architecture. We demonstrate this for auction (V) with a training set of 20 samples and plot the training and test losses as a function of time (measures in epochs) for both architectures in Figure 4.

Out-of-setting generalization. The number of parameters in our permutation-equivariant architecture is independent from the input's size. Given an architecture that was trained on samples of size (n, m), it is possible to evaluate it on samples of any size (n', m') (More details in Appendix A). This evaluation is not well defined for feed-forward architectures where the dimension of the weights depends on the input size. We use this advantage to check whether models trained in a fixed setting perform well in totally different ones.

(α) Train an equivariant architecture on 1 bidder, 5 items and test it on 1 bidder, n items for $n = 2 \cdots 10$. All the items values are sampled independently from $\mathcal{U}[0, 1]$.

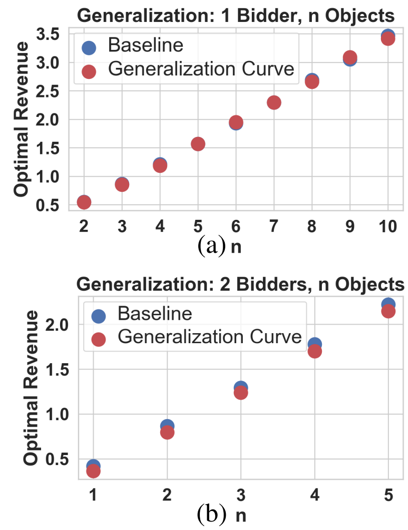

Figure 5:Generalization revenue of EquivariantNet in experiment (α) and (β). Each baseline point is computed using a RegretNet architecture trained from scratch.

(β) Train an equivariant architecture on 2 bidders, 3 objects and test it on 2 bidders, n objects for $n = 2 \cdots 6$. All the items values are sampled independently from $\mathcal{U}[0, 1]$.

Figure 5 (a)-(b) reports the test revenue that we get for different values of n in (α) and (β) and compares it to the empirical optimal revenue. Surprisingly, our model does generalize well. It is worth mentioning that knowing how to solve a larger problem such as 1×5 does not automatically result in a capacity to solve a smaller one such as 1×2; the generalization does happen on both ends. Our approach looks promising regarding out of setting generalization. It generalizes well when the number of objects varies and the number of bidders remain constants. However, generalization to settings where the number of bidders varies is more difficult due to the complex interactions between bidders. We do not observe good generalization with our current method.

Conclusion

We explored the effect of adding domain knowledge in neural network architectures for auction design. We built a permutation-equivariant architecture to design symmetric auctions and highlighted its multiple advantages. It recovers several known optimal results and provides competitive results in asymmetric auctions. Compared to fully connected architectures, it is more sample efficient and is able to generalize to settings it was not trained on.

Our architecture presents some limitations. It assumes that all the bidders and items are permutation-equivariant. However, in some real-world auctions, the item/bidder-symmetry only holds for a group of bidders/items. More advanced architectures such as Equivariant Graph Networks (Maron et al. 2018) may solve this issue. Another limitation is that we only consider additive valuations. An interesting direction would be to extend our approach to other settings as unit-demand or combinatorial auctions.

References

Alaei, S. 2011. Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers. In *the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Alaei, S.; Fu, H.; Haghpanah, N.; and Hartline, J. 2013. The Simple Economics of Approximately Optimal Auctions. In *the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Alaei, S.; Fu, H.; Haghpanah, N.; Hartline, J.; and Malekian, A. 2012. Bayesian Optimal Auctions via Multi- to Single-agent Reduction. In *the 13th ACM Conference on Electronic Commerce (EC)*.

Alaei, S.; Hartline, J. D.; Niazadeh, R.; Pountourakis, E.; and Yuan, Y. 2015. Optimal Auctions vs. Anonymous Pricing. *CoRR* abs/1507.02615. URL http://arxiv.org/abs/1507.02615.

Babaioff, M.; Immorlica, N.; Lucier, B.; and Weinberg, S. M. 2014. A simple and approximately optimal mechanism for an additive buyer. *SIGecom Exchanges* 13(2): 31–35. doi:10.1145/2728732.2728736. URL http://doi.acm.org/10.1145/2728732.2728736.

Balcan, M.-F. F.; Sandholm, T.; and Vitercik, E. 2016. Sample complexity of automated mechanism design. In *Advances in Neural Information Processing Systems*, 2083–2091.

Beyhaghi, H.; and Weinberg, S. M. 2019. Optimal (and Benchmark-Optimal) Competition Complexity for Additive Buyers over Independent Items. In *Proceedings of the 51st ACM Symposium on Theory of Computing Conference (STOC)*.

Cai, Y.; Daskalakis, C.; and Weinberg, S. M. 2012a. An algorithmic characterization of multi-dimensional mechanisms. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, 459–478. doi:10.1145/2213977.2214021. URL https://doi.org/10.1145/2213977.2214021.

Cai, Y.; Daskalakis, C.; and Weinberg, S. M. 2012b. Optimal Multi-dimensional Mechanism Design: Reducing Revenue to Welfare Maximization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, 130–139.

Cole, R.; and Roughgarden, T. 2014. The sample complexity of revenue maximization. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, 243–252.

Conitzer, V.; and Sandholm, T. 2002. Complexity of mechanism design. *arXiv preprint cs/0205075* .

Conitzer, V.; and Sandholm, T. 2004. Self-interested automated mechanism design and implications for optimal combinatorial auctions. In *Proceedings of the 5th ACM conference on Electronic commerce*, 132–141.

Daskalakis, C.; Deckelbaum, A.; and Tzamos, C. 2017. Strong Duality for a Multiple-Good Monopolist. *Econometrica* 85(3): 735–767.

Daskalakis, C.; and Weinberg, S. M. 2012. Symmetries and optimal multi-dimensional mechanism design. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 370–387.

Dhangwatnotai, P.; Roughgarden, T.; and Yan, Q. 2015. Revenue maximization with a single sample. *Games and Economic Behavior* 91: 318–333. doi:10.1016/j.geb.2014.03.011. URL https://doi.org/10.1016/j.geb.2014.03.011.

Duetting, P.; Feng, Z.; Narasimhan, H.; Parkes, D.; and Ravindranath, S. S. 2019. Optimal Auctions through Deep Learning. In Chaudhuri, K.; and Salakhutdinov, R., eds., *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, 1706–1715. Long Beach, California, USA: PMLR. URL http://proceedings.mlr.press/v97/duetting19a.html.

Dughmi, S.; Han, L.; and Nisan, N. 2014. Sampling and representation complexity of revenue maximization. In *International Conference on Web and Internet Economics*, 277–291. Springer.

Dütting, P.; Fischer, F.; Jirapinyo, P.; Lai, J. K.; Lubin, B.; and Parkes, D. C. 2015. Payment rules through discriminant-based classifiers. *ACM Transactions on Economics and Computation (TEAC)* 3(1): 1–41.

Feng, Z.; Narasimhan, H.; and Parkes, D. C. 2018. Deep learning for revenue-optimal auctions with budgets. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems*, 354–362. International Foundation for Autonomous Agents and Multiagent Systems.

Golowich, N.; Narasimhan, H.; and Parkes, D. C. 2018. Deep Learning for Multi-Facility Location Mechanism Design. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI 2018)*, 261–267. URL https://econcs.seas.harvard.edu/files/econcs/files/golowich_ijcai18.pdf.

Gonczarowski, Y. A.; and Nisan, N. 2017. Efficient empirical revenue maximization in single-parameter auction environments. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, 856–868. doi:10.1145/3055399.3055427. URL http://doi.acm.org/10.1145/3055399.3055427.

Gonczarowski, Y. A.; and Weinberg, S. M. 2018. The Sample Complexity of up-to-ε multidimensional revenue maximization. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS*.

Guo, C.; Huang, Z.; and Zhang, X. 2019. Settling the sample complexity of single-parameter revenue maximization. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019.*, 662–673. doi:10.1145/3313276.3316325. URL https://doi.org/10.1145/3313276.3316325.

Guo, M.; and Conitzer, V. 2010. Computationally feasible automated mechanism design: General approach and case studies. In *Twenty-Fourth AAAI Conference on Artificial Intelligence*.

Hart, S.; and Nisan, N. 2012. Approximate Revenue Maximization with Multiple Items. In *the 13th ACM Conference on Electronic Commerce (EC)*.

Hart, S.; and Nisan, N. 2013. The menu-size complexity of auctions. In *the 14th ACM Conference on Electronic Commerce (EC)*.

Hart, S.; and Reny, P. J. 2015. Maximizing Revenue with Multiple Goods: Nonmonotonicity and Other Observations. *Theoretical Economics* 10(3): 893–922.

Hartford, J.; Graham, D. R.; Leyton-Brown, K.; and Ravanbakhsh, S. 2018. Deep models of interactions across sets. *arXiv preprint arXiv:1803.02879* .

Hartline, J. D.; and Roughgarden, T. 2009. Simple versus optimal mechanisms. In *ACM Conference on Electronic Commerce*, 225–234.

Hartline, J. D.; and Taggart, S. 2019. Sample Complexity for Non-Truthful Mechanisms. In *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019.*, 399–416. doi:10.1145/3328526.3329632. URL https://doi.org/10.1145/3328526.3329632.

Huang, Z.; Mansour, Y.; and Roughgarden, T. 2018. Making the most of your samples. *SIAM Journal on Computing* 47(3): 651–674.

Jin, Y.; Lu, P.; Qi, Q.; Tang, Z. G.; and Xiao, T. 2019a. Tight approximation ratio of anonymous pricing. In Charikar, M.; and Cohen, E., eds., *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, 674–685. ACM. doi:10.1145/3313276.3316331. URL https://doi.org/10.1145/3313276.3316331.

Jin, Y.; Lu, P.; Tang, Z. G.; and Xiao, T. 2019b. Tight revenue gaps among simple mechanisms. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 209–228. SIAM.

Kothari, P.; Mohan, D.; Schvartzman, A.; Singla, S.; and Weinberg, S. M. 2019a. Approximation schemes for a buyer with independent items via symmetries. *arXiv preprint arXiv:1905.05231* .

Kothari, P.; Singla, S.; Mohan, D.; Schvartzman, A.; and Weinberg, S. M. 2019b. Approximation Schemes for a Unit-Demand Buyer with Independent Items via Symmetries. *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)* doi:10.1109/focs.2019.00023.

Lahaie, S. 2011. A kernel-based iterative combinatorial auction. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*.

Li, X.; and Yao, A. C.-C. 2013. On Revenue Maximization for Selling Multiple Independently Distributed Items. *Proceedings of the National Academy of Sciences* 110(28): 11232–11237.

Manelli, A.; and Vincent, D. 2006. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory* 127(1): 1–35.

Manelli, A. M.; and Vincent, D. R. 2010. Bayesian and Dominant-Strategy Implementation in the Independent Private-Values Model. *Econometrica* 78(6): 1905–1938.

Maron, H.; Ben-Hamu, H.; Shamir, N.; and Lipman, Y. 2018. Invariant and equivariant graph networks. *arXiv preprint arXiv:1812.09902* .

Medina, A. M.; and Mohri, M. 2014. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, 262–270.

Morgenstern, J.; and Roughgarden, T. 2016. Learning simple auctions. In *Conference on Learning Theory*, 1298–1318.

Morgenstern, J. H.; and Roughgarden, T. 2015. On the pseudo-dimension of nearly optimal auctions. In *Advances in Neural Information Processing Systems*, 136–144.

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

Narasimhan, H.; and Parkes, D. C. 2016. A general statistical framework for designing strategy-proof assignment mechanisms. In *UAI'16 Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*.

Pavlov, G. 2011. Optimal Mechanism for Selling Two Goods. *The B.E. Journal of Theoretical Economics* 11(3).

Rahme, J.; Jelassi, S.; and Weinberg, S. M. 2021. Auction Learning as a Two-Player Game. In *International Conference on Learning Representations*. URL https://openreview.net/forum?id=YHdeAO61l6T.

Sandholm, T.; and Likhodedov, A. 2015. Automated design of revenue-maximizing combinatorial auctions. *Operations Research* 63(5): 1000–1025.

Shen, W.; Tang, P.; and Zuo, S. 2019. Automated mechanism design via neural networks. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, 215–223. International Foundation for Autonomous Agents and Multiagent Systems.

Tacchetti, A.; Strouse, D.; Garnelo, M.; Graepel, T.; and Bachrach, Y. 2019. A Neural Architecture for Designing Truthful and Efficient Auctions. *arXiv preprint arXiv:1907.05181* .

Vickrey, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance* 16(1): 8–37.

Wang, Z.; and Tang, P. 2014. Optimal mechanisms with simple menus. In *ACM Conference on Economics and Computation, EC '14, Stanford , CA, USA, June 8-12, 2014*, 227–240. doi:10.1145/2600057.2602863. URL https://doi.org/10.1145/2600057.2602863.