Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

LEARNING PAYMENT-FREE RESOURCE ALLOCATION MECHANISMS

Sihan Zeng[1], Sujay Bhatt[1], Eleonora Kreacic[2], Parisa Hassanzadeh[3], Alec Koppel[1], and Sumitra Ganesh[1]

[1]J.P. Morgan AI Research, New York, NY, USA
[2]J.P. Morgan AI Research, London, UK
[3]Samsung Data Systems, San Jose, CA, USA

ABSTRACT

We consider the design of mechanisms that allocate limited resources among self-interested agents using neural networks. Unlike the recent works that leverage machine learning for revenue maximization in auctions, we consider welfare maximization as the key objective in the *payment-free* setting. Without payment exchange, it is unclear how we can align agents' incentives to achieve the desired objectives of truthfulness and social welfare simultaneously, without resorting to approximations. Our work makes novel contributions by designing an *approximate* mechanism that desirably trade-off social welfare with truthfulness. Specifically, (i) we contribute a new end-to-end neural network architecture, `ExS-Net`, that accommodates the idea of "money-burning" for mechanism design without payments; (ii) we provide a generalization bound that guarantees the mechanism performance when trained under finite samples; and (iii) we provide an experimental demonstration of the merits of the proposed mechanism.

1 INTRODUCTION

Mechanism design studies how to induce a game among strategic agents in such a way that the induced game satisfies a set of desired properties at the equilibrium. These properties include individually rationality (agents are motivated to participate in the game), incentive compatibility (agents are motivated to report private information truthfully), social welfare (socially desirable outcome is chosen), efficiency (resources are not wasted when there is demand), envy-freeness (agents do not wish for others' share of the resources); to name a few. In this work, we are interested in *incentive compatibility*, *social welfare* and *efficiency*; and we propose learning the rules of the game so that the agents are motivated to report their preferences truthfully in a *payment-free setting*, and a socially desirable outcome that is fair and efficient results in the equilibrium.

1.1 Main Contribution

We consider multiple-agents and multiple divisible items and aim to design an incentive compatible (**IC**) mechanism that maximizes social welfare & efficiency in the payments-free setting. Since the three objectives cannot be achieved simultaneously as shown in Cole et al. (2013), we consider learning an approximate mechanism that achieve a desirable trade-off. To provide an overview of our work:

- We consider Nash Social Welfare (NSW) as the welfare objective for the mechanism designer. We further impose constraints on the approximate incentive compatibility of the mechanism. These constraints are defined using the notion of "exploitability" (Goktas and Greenwald 2022) – which measures the maximum utility gain when agents deviate from truthfully reporting.
- We design a novel "state-augmentation with an artificial agent" based neural architecture – named `ExS-Net` – that simulates "money-burning" (Hartline and Roughgarden 2008) in the "hardware", i.e., intentional withholding of resources as an implicit form of payment. In addition, we make the

design modular, in that, any neural network (feed-forward, CNN, Transformers) can be used as the hidden layer and the money-burning hardware only modifies the output/ activation layer.
- We derive the generalization bounds for `Ex-Net` and also provide guarantees under distribution shift. The robustness to distribution shift in the experiments, specifically, demonstrates the significance of our contribution by allowing the training data itself to be subjected to adversarial contamination.
- Extensive experiments confirm the efficacy of the proposed mechanisms in achieving the desired trade-off. We explicitly evaluate the exploitability and social welfare of our mechanism compared to the baseline mechanisms Proportionally Fair (PF) and Partial Allocation (PA) (Cole et al. 2013). The proposed mechanism `ExS-Net` significantly outperforms PF in terms of exploitability and PA in terms of efficiency & social welfare.

1.2 Motivating Examples

We include a few potential applications that motivate the setting considered in this paper. Key aspects include welfare maximizing, efficient, and IC resource allocation without the payments.

- **Computing**: Firm allocates budget constrained computing resources (memory, GPUs) to various internal business and research teams for assisting in their individual goals. Each team is self-interested and could potentially misreport the preferences and demand to gain increased allocations. A common occurrence is when the research team is facing an impending conference deadline. Designing fair allocation mechanisms that ensures that the teams have little incentive to report untruthfully and are efficient, are necessary for cost reduction and improving overall satisfaction.
- **Auto-loans**: Self-interested customers request for loan packages from the firm via dealers and in-turn share experiences on third-party trusted websites. Dealers act as intermediaries for vetting the customer application and bringing in the business. Designing mechanisms that secure fair & efficient loan approvals for customers while ensuring they share truthful account of their experiences on the platform is important in achieving AI for social good.

1.3 Related Literature

We organize the literature review into the different resource allocation problem classes, relevant solution approaches, and the key differences from the most relevant work.

Resource Allocation Settings

Classic work on resource allocation focuses on divisible resources, with recent interest spanning both the indivisible and divisible setting. In either case, payment is a frequently used tool to align incentives (Pavlov 2011; Giannakopoulos and Koutsoupias 2014; Dütting et al. 2024) and ensure the truthful behavior of resource consumers. However, payment is naturally forbidden in many settings including organ donation, food and necessity distribution by charity, allocation of GPU hours by an institution to its employees (Dekel et al. 2010; Procaccia and Tennenholtz 2013). Our work focuses on the divisible setting with multiple-agents and items, and considers designing a fair & efficient resource allocation mechanism that is incentive compatible without payments. Since the three objectives cannot be achieved simultaneously, we consider learning approximate mechanisms that achieves a desirable trade-off.

Solution Approaches for Payment-Free Setting

Ensuring allocation is fair among agents may be formalized through *proportional fairness* (**PF**) (Kelly 1997). PF allocations are such that if an alternate allocation is adopted, the percentage utility gain over all agents sums up to a non-positive number. Importantly, PF is achieved by maximizing Nash social welfare (**NSW**), the product of all agents' utilities (Bertsimas et al. 2011), which we use as a quantifier of social welfare in this work. The PF mechanism is one that directly seeks to optimize NSW by solving a constrained optimization program, and by definition achieves the maximum possible NSW under allocation constraints.

Cole et al. (2013) identifies a key issue with the PF mechanism – strategic agents may misreport their utility functions to gain inflated allocation from a supplier running the PF mechanism. To prevent such untruthful behavior, Cole et al. (2013) design a novel mechanism – partial allocation (**PA**) – which achieves the exact IC at the cost of efficiency and NSW. Prior to our work, PA is the state-of-the-art in balancing NSW and incentive compatibility. The main contribution of our work is to achieve an even more desirable operating point between NSW and incentive compatibility.

Learning-Based Mechanism Design in Resource Allocation

In the literature on mechanism design for resource allocation, key gaps remain due to the difficulty of hand-designing approximate mechanisms that are "somewhere in-between" the desired objectives. For example, the PF mechanism optimizes NSW but is not IC; PA is hand-design to achieve IC but suffers sub-optimal NSW. Recently, there has been efforts to take a data-driven approach to resource allocation using neural networks. Dütting et al. (2024) leads the way by learning payment-based allocation mechanism that not only approximate well-known optimal solutions for special cases, but also demonstrate how the solutions look like for multi-item & multi-agent settings where the optimal solution is unknown. Our work deviates from the existing literature as we have to design IC mechanisms in the more challenging payment-free setting. The only known approach to align the incentive of the agents with that of the supplier without payment is "money-burning" (Hartline and Roughgarden 2008) – intentional withholding of resources as an implicit form of payment. Inspired by this domain knowledge, we design a neural-network-based mechanism that approximates "money-burning".

2 MECHANISM DESIGN WITHOUT PAYMENTS

A supplier allocates a finite number M of divisible resources to N agents. The allocation can be represented as a vector $a \in \mathbb{R}^{NM}$, where $a_{i,m}$ represents the quantity of resource $m \in [M]$ allocated to agent $n \in [N]$. The supplier observes the budget $b_m \geq 0$ on each resource m, which specifies the quantity of the resource available. We denote $b = [b_1, \cdots, b_M]^\top \in \mathscr{B} \subseteq \mathbb{R}_+^M$, where $\mathscr{B}$ is the set of all possible budgets. The agent's satisfaction with an allocation outcome is measured by a utility function. In this work, we assume that every agent has a (threshold) additive linear utility – each unit of resource m increases the utility of agent i by value $v_{i,m}$ up to the threshold $x_{i,m}$ which represents demand. We aggregate valuations and demands by defining $v_i = [v_{i,1}, \cdots, v_{i,M}]^\top$, $x_i = [x_{i,1}, \cdots, x_{i,M}]^\top$, and $v = [v_1^\top, \cdots, v_N^\top]^\top$, $x = [x_1^\top, \cdots, x_N^\top]^\top$. These quantities are in respective spaces of values and demand $\mathscr{V} \subseteq [\underline{v}, \overline{v}]^{NM}$, $\mathscr{D} \subseteq [\underline{d}, \overline{d}]^{NM}$ for some scalars $\underline{v}, \overline{v} > 0$, $\underline{d}, \overline{d} \geq 0$. Given an allocation outcome $a \in \mathbb{R}^{NM}$, the utility function of agent i is $u_i : \mathbb{R}_+^{NM} \times \mathscr{V} \times \mathscr{D} \to \mathbb{R}_+$ expressed as

$$u_i(a, v, x) \triangleq \sum_{m=1}^{M} v_{i,m} \min\{a_{i,m}, x_{i,m}\}. \tag{1}$$

This utility models satisfaction in real-life problems and is widely studied (Bliem et al. 2016; Camacho et al. 2021). Our work operates in the common setting where the supplier knows the functional form of the utility function and relies on each agent i to report its parameters $v_{i,m}$ and $x_{i,m}$ (Procaccia and Tennenholtz 2013; Cole et al. 2013; Dütting et al. 2024).

We say a mapping $f : \mathscr{V} \times \mathscr{D} \times \mathscr{B} \to \mathbb{R}_+^{NM}$ is an allocation mechanism, i.e, a mechanism computes the allocations given the values and demands of the agents.

We consider Nash Social Welfare (NSW) as the welfare objective for the mechanism designer, which balances social welfare and efficiency. Unlike the egalitarian notion that aims to maximize the utility of the least satisfied agent irrespective of the inefficiency and the utilitarian notion of welfare that maximizes efficiency while disregarding the welfare, NSW sits in-between (Bertsimas et al. 2011) and achieves a good trade-off.

A mechanism may further encode agent priorities via a weights vector $w \in \mathbb{R}_+^N$, that are determined solely by the supplier before the agents reveal their requests.

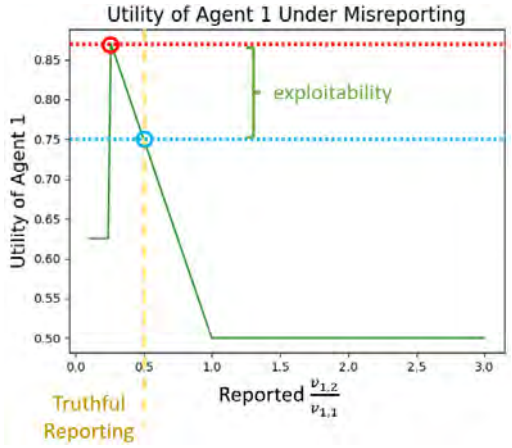

Figure 1: The utility of agent 1 as it varies the reported preference ratio $v_{1,2}/v_{1,1}$ from 0.1 to 3. Blue line indicates the utility of agent 1 under truthful report, and red line indicates the maximum achievable utility under misreport. Exploitability of PF is shown as their gap. Agent 1 increases its utility by under-reporting $v_{1,2}/v_{1,1}$.

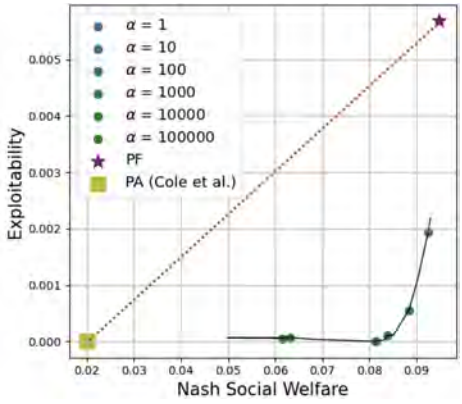

Figure 2: Performance trade-off frontier. The solid curve obtained by proposed mechanism `ExS-Net` over a range of exploitability tolerance (see Definition 3). The figure illustrates the superior exploitability/NSW trade-off achieved by `ExS-Net` over the mixture of PA and PF indicated by the dotted line (details in Section 5.4).

Definition 1 (Nash Social Welfare)**.** Given any mechanism f and $v \in \mathscr{V}, x \in \mathscr{D}, b \in \mathscr{B}$, NSW is defined as

$$\begin{aligned} \mathbf{NSW}(f,v,x,b) &\triangleq \Pi_{i=1}^{N} u_i(f,v,x)^{w_i}, \\ \mathbf{logNSW}(f,v,x,b) &\triangleq \log\Big(\mathbf{NSW}(f,v,x,b)\Big) = \textstyle\sum_{i=1}^{N} w_i \log(u_i(f,v,x)) \ . \end{aligned} \tag{2}$$

We next formalize *exploitability* as the maximum gain an agent may obtain by misreporting its utility function.

Definition 2 (Exploitability)**.** Under mechanism f and $v \in \mathscr{V}, x \in \mathscr{D}, b \in \mathscr{B}$, we define the exploitability at agent i as its largest possible utility increase due to misreporting, given that the every other agent reports its true valuation and demand, i.e.,

$$\mathbf{expl}_i(f,v,x,b) \triangleq \max_{v_i', x_i'} u_i\big(f((v_i', v_{-i}), (x_i', x_{-i}), b), v, x\big) - u_i\big(f(v,x,b), v, x\big).$$

A mechanism f is ***incentive compatible*** (IC) if $\forall\ i, v, x, b$ $\mathbf{expl}_i(f,v,x,b) = 0$.

The goal in incentive compatible mechanism design to promote social welfare while incurring zero or low exploitability for all agents. It is well-known that no mechanism can simultaneously achieve the optimal NSW and perfect incentive compatibility (Cole et al. 2013). In this work, we aim to achieve a desirable trade-off among the two metrics.

2.1 State-of-the-art Mechanisms

1. *Proportional Fairness mechanism* (f^{PF}): This mechanism was first introduced in Kelly (1997). Assuming that the agents *truthfully* report their preferences, f^{PF} achieves the optimal NSW (Bertsimas et al. 2011), but the constraint on exploitability is violated even in the simplest instances. As an example, we show how utility gain may arise from misreporting under the PF mechanism in a two-agent two-resource system. When agent 2 reports truthfully, Figure 1 plots the utility of agent 1 [cf. (1)] as it varies the reported preference ratio $v_{1,2}/v_{1,1}$ from 0.1 to 3 (the true ratio is 0.5). With

the dashed line indicating the utility of agent 1 under truthful reporting, under-reporting $v_{1,2}/v_{1,1}$ increases agent 1's utility up to 17%, which can be a huge incentive in practical applications.

Mechanism: Proportional Fairness

$$f^{PF}(v,x,b) = \underset{a\in\mathbb{R}^{NM}}{\operatorname{argmin}} -\sum_{i=1}^{N} w_i \log(a_i^\top v_i)$$
$$\text{s.t. } 0 \le a \le x,\ \sum_{i=1}^{N} a_{i,m} \le b_m, \forall m.$$

2. *Partial Allocation mechanism* (f^{PA}): Proposed in Cole et al. (2013), this mechanism is built upon the PF mechanism and withholds resources according to an externality ratio. The PA mechanism achieves *zero* exploitability, but each agent can only be guaranteed to receive a $1/e$ fraction of the resources that it would receive under the PF mechanism, meaning that there is resource waste and a significant reduction in NSW. Let $x^{-i} \in \mathbb{R}^{N\times M}$ denote a modified demand matrix such that $x_j^{-i} = x_j \in \mathbb{R}^M$ for all $j \neq i$ and $x_i^{-i} = 0$. We summarize the steps of the PA mechanism as follows.

Mechanism: Partial Allocation

(1) For each agent i, calculate $a^{-i,\star}$, the PF allocation outcome that would arise in the absence of agent i

$$a^{-i,\star} \triangleq f^{PF}(v,x^{-i},b).$$

(2) Compute the externality ratio $r_i \in \mathbb{R}_+$ for agent i as

$$r_i = \left(\frac{\prod_{j\neq i} u_j(f^{PF}(v,x,b),v,x)^{w_j}}{\prod_{j\neq i} u_j(a^{-i,\star},v,x)^{w_j}}\right)^{1/w_i}.$$

(3) Allocate $f_i^{PA}(v,x,b) \in \mathbb{R}^M$ to agent i according to

$$f_i^{PA}(v,x,b) = r_i f_i^{PF}(v,x,b).$$

No mechanisms in the literature balance these criteria in the payment-free setting when agents may misreport: PA mechanism is sub-optimal in NSW, whereas PF is optimal in NSW but incurs high exploitability. This dichotomy does not exist in the *auction setting* as payments are used as an enforcement tool for truthful reporting. Classic auction mechanisms including VCG mechanism and Myerson mechanism balance IC with revenue maximization (Nisan 2007).

The key challenge that explains the gap in the literature is the difficulty to hand-design mechanisms that trade-off conflicting objectives – it is easier to design those that satisfy some of them *exactly*. This motivates the work in this paper: We propose a learning based approach to mechanism design that achieves a desirable trade-off between the objectives, in the payment-free setting. As illustrated in Figure 2, by adjusting a tunable weight parameter, we learn a mechanism (`ExS-Net`) that achieves a NSW-exploitability trade-off frontier that significantly improves over PF and PA. Simulation details will be discussed later in Section 5.4.

3 LEARNING PAYMENT-FREE MECHANISM

In this section, we design a novel "state-augmentation with an artificial agent" based neural architecture that simulates "money-burning" (Hartline and Roughgarden 2008) in the "hardware", i.e., intentionally withholding of resources as an implicit form of payment.

3.1 `ExS-Net` – Strategic Resource Withholding

We name the mechanism `ExS-Net`, which stands for **Ex**ploitability-Aware Network with **S**oftmax activation, and provide a schematic representation in Figure 3. The `ExS-Net` mechanism can be represented as a

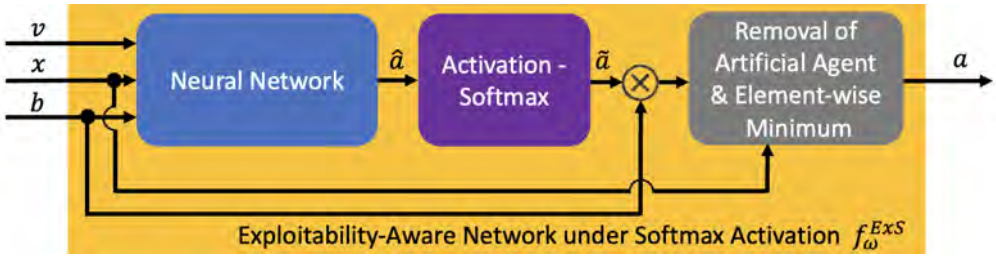

Figure 3: Exs-Net pipeline.

function $f^\omega : \mathscr{V} \times \mathscr{D} \times \mathscr{B} \to \mathbb{R}_+^{NM}$ parameterized by neural network parameter ω. Specifically, we employ a neural network that maps v,x,b to an initial allocation weight vector $\hat{a} \in \mathbb{R}^{(N+1)M}$ (rather than $\mathbb{R}^{NM}$). As pointed out in Hartline and Roughgarden (2008), IC cannot be achieved in the payment-free setting without resource withholding ("money burning"). Motivated by this knowledge, we expand the space of allocations to introduce a synthetic agent that receives the proportion of the allocation that will eventually be withheld. This synthetic agent enlarges the allocation dimension in $\hat{a}$. Through the use of a softmax function applied to $\hat{a}$ along the direction of agents, ExS-Net calculates the allocation ratio $\tilde{a} \in \mathbb{R}^{(N+1)M}$ such that $\sum_{i=1}^{N+1} \tilde{a}_{i,m} = 1$, i.e.,

$$\tilde{a}_{i,m} = \exp(\hat{a}_{i,m}) / \left(\sum_{i'=1}^{N+1} \exp(\hat{a}_{i',m})\right) \text{ for all } m.$$

As the fraction of resource m assigned to agent i, the quantity $\tilde{a}_{i,m}$ leads to allocation $a_{i,m} = \min\{\tilde{a}_{i,m} b_m, x_{i,m}\}$. We note that our design is modular in that any neural network (feed-forward, CNN, Transformers) can be used as the hidden layer and the money-burning hardware only modifies the output/activation layer.

3.2 Training `ExS-Net` Mechanism

Assuming that the values and demands of the agents and the budgets on the resources follow a joint distribution F, we use samples from F to train a mechanism that optimize NSW while staying approximately incentive compatible in expectation. To this end, we define the notion of ε-incentive compatibility in expectation.

Definition 3. [ε-Incentive Compatibility] A mechanism f is ε-incentive compatible over distribution F if

$$\mathbb{E}_{(v,x,b)\sim F}[\mathbf{expl}_i(f,v,x,b)] \leq \varepsilon, \quad \forall i. \tag{3}$$

We define learning ω as the maximization of expected NSW with an ε-incentive compatibility constraint. The learning objective is given as follows.

Learning Objective

$$\max_\omega \mathbb{E}_{(v,x,b)\sim F}[\mathbf{logNSW}(f^\omega,v,x,b)] \quad \text{s.t. } \mathbb{E}_{(v,x,b)\sim F}[\mathbf{expl}_i(f^\omega,v,x,b)] \leq \varepsilon, \forall i. \tag{4}$$

In practice, one may not have access to the distribution F, but instead a training set of values, demand, and budgets $\{(v^l,x^l,b^l) \sim F\}_{l=1}^L$. We train the mechanisms with the finite dataset via empirical risk minimization (ERM) by forming the sample-averaged estimates of the expected NSW and exploitability.

$$\max_\omega \sum_{l=1}^L \mathbf{logNSW}(f^\omega,v^l,x^l,b^l) \quad \text{s.t. } \sum_{l=1}^L \mathbf{expl}_i(f^\omega,v^l,x^l,b^l) \leq \varepsilon, \forall i. \tag{5}$$

We note that (5) deviates from the conventional supervised learning objectives – we do not need paired ground truth data. Instead, the objective (5) can be evaluated and optimized with a dataset on truthfully reported valuations and demands only, which can be elicited under an existing incentive compatible mechanism such as the PA mechanism.

Algorithm 1: Training ExS-Net

Input: Initial network parameter $\omega^{[0]}$, dual variables $\{\gamma_i^{[0]}\}_{i=1}^N$, training dataset $\{(v^l, x^l, b^l)\}_{l=1}^L$, batch size s, training iterations K, primal and dual learning rate α, β

Output: Network parameter $\omega^{[K]}$.

for $k = 0, 1, \cdots, K-1$ **do**

1) Randomly draw sample index set $\mathscr{S}^{[k]}$ with $|\mathscr{S}^{[k]}| = s$ and compute empirical logNSW and exploitability

$$\widehat{\mathbf{logNSW}}^{[k]} = \sum_{l \in \mathscr{S}^{[k]}} \mathbf{logNSW}(f^{\omega^{[k]}}, v^l, x^l, b^l),$$
$$\widehat{\mathbf{expl}}_i^{[k]} = \sum_{l \in \mathscr{S}^{[k]}} \mathbf{expl}_i(f^{\omega^{[k]}}, v^l, x^l, b^l), \quad \forall i$$

2) Neural network parameter update:

$$\omega^{[k+1]} = \omega^{[k]} - \alpha \nabla_{\omega^{[k]}} \Big(\sum_i \gamma_i^{[k]} \widehat{\mathbf{expl}}_i^{[k]} - \widehat{\mathbf{logNSW}}^{[k]} \Big)$$

3) Dual variable update:

$$\gamma_i^{[k+1]} = \Pi_+ \Big(\gamma_i^{[k]} + \beta \big(\widehat{\mathbf{expl}}_i^{[k]} - \varepsilon \big) \Big), \quad \forall i$$

end

We optimize (5) with respect to ω by using a simple primal-dual gradient descent-ascent algorithm to find the saddle point of the Lagrangian. We present the training scheme in Alg. 1, where Π_+ denotes the projection of a scalar to the non-negative range. The dual variable $\gamma_i \in \mathbb{R}_+$ is associated with the i_{th} exploitability constraint in (5). In essense, the neural network parameter is updated to optimize the sum of log NSW and exploitability, with the weight of exploitability adjusted dynamically according to the level of constraint violation.

3.3 Inference Using `ExS-Net`

We use $\omega^\star$ to denote the neural network parameter returned in the training procedure. In the inference phase, we need to generate allocations given budgets and agent-reported demands & valuations using `ExS-Net` parameterized by $\omega^\star$.

Mechanism: `ExS-Net`

(1) Given input budget b, reported valuation v, and reported demands x, compute $\hat{a} \in \mathbb{R}^{(N+1)M}$ as output of neural network parameterized by $\omega^\star$

(2) Soft-max function

$$\tilde{a}_{i,m} = \exp(\hat{a}_{i,m}) / \big(\textstyle\sum_{i'=1}^{N+1} \exp(\hat{a}_{i',m}) \big) \text{ for all } m.$$

3) Determine allocation $f^{\omega^\star}(v, x, b) = a \in \mathbb{R}^{NM}$ such that

$$a_{i,m} = \min\{\tilde{a}_{i,m} b_m, x_{i,m}\} \text{ for all } i = 1, \cdots, N, m = 1, \cdots, M.$$

Discard remaining resources.

We note that while ExS-Net is given the ability to discard resources with the aim of achieving approximate incentive compatibility, we observe through numerical simulations that the quantity it learns to discard is small. This is reflected by the near-optimal efficiency and high NSW of the mechanism showcased in Section 5. In other words, we can reduce the exploitability (in the experiments over 80% reduction is observed on average) at the cost of a small compromise in efficiency and NSW.

4 GENERALIZATION BOUNDS

We establish the convergence of the learned mechanisms by bounding the sub-sampling error by a sub-linear function of the batch size, which matches recent rates for auctions (Dütting et al. 2024) and ensures that the objective of (5) converges to (4) with sufficient samples.

For mechanism f, we define the generalization errors with L samples as

$$\varepsilon_{\textbf{logNSW}}(f,L) = -\mathbb{E}_{(v,x,b)\sim F}[\textbf{logNSW}(f,v,x,b)] + \textstyle\sum_{l=1}^{L} \textbf{logNSW}(f,v^l,x^l,b^l),$$
$$\varepsilon_{\textbf{expl},i}(f,L) = \mathbb{E}_{(v,x,b)\sim F}[\textbf{expl}_i(f,v,x,b)] - \textstyle\sum_{l=1}^{L} \textbf{expl}_i(f,v^l,x^l,b^l).$$

Theorem 1 (Generalization Bound)**.** Consider ExS-Net f^{ω} mechanisms parameterized by a neural network with R hidden layers, K nodes per hidden layer, ReLU activation, a total of d parameters, and the vector of all model parameters $\|\omega\|_1 \leq \Omega$. With probability at least $1-\delta$,

$$\max\{\varepsilon_{\textbf{logNSW}}(f^{\omega},L), \varepsilon_{\textbf{expl},i}(f^{\omega},L)\} \leq \mathcal{O}\Big(\frac{\sqrt{Rd\log(LN\Omega\max\{K,MN\})}}{L} + N\sqrt{\frac{\log(1/\delta)}{L}}\Big).$$

The result provides that the generalization error decays at rate $\mathcal{O}(L^{-1/2})$ where L is the training sample size. In auction design with payment, RegretNet, the state-of-the-art learned mechanism proposed in Dütting et al. (2024), also achieves a $\mathcal{O}(L^{-1/2})$ rate, which we match in the non-payment setting.

5 EXPERIMENTS

In this section, we demonstrate the effectiveness and robustness of the proposed ExS-Net mechanism from various perspectives. First, we evaluate the social welfare, efficiency, and exploitability of the mechanism in systems with different numbers of agents and resources. Second, we study distribution mismatch, a situation where the mechanism is required to perform under a test distribution different from that observed during training. Finally, we adjust the parameter ε in (5), which controls the trade-off between social welfare and exploitability, and show the superior trade-off frontier.

5.1 Simulation Setup

We use a four-layer neural network as the function approximation for ExS-Net in this work. The specific simulation setup including data distribution, baselines, and evaluation metrics are discussed below.

Data Generation. In all experiments, values and demands follow uniform and Bernoulli uniform distributions, respectively, within the range $[0.1,1]$. Specifically, we sample

$$v_{i,m} \sim \text{Unif}(0.1,1), \quad \widetilde{x}_{i,m} \sim \text{Unif}(0.1,1), \quad \widehat{x}_{i,m} \sim \text{Bern}(0.5), \quad x_{i,m} = \widetilde{x}_{i,m}\widehat{x}_{i,m}. \tag{6}$$

We make $x_{i,m}$ a product of Bernoulli and uniform random variables as we would like to represent the case where not all agents request all resources. Unless noted otherwise, the budget for every resource is set to $\frac{N}{2}$, for number of agents N. This creates competition for resources in expectation. The agent weights are set to 1. Test data is sampled from the same distributions as training data except in Section 5.3.

Baselines. We evaluate against the PF & PA mechanisms. Both PF and PA are hand-designed mechanisms that do not require training, but are time-consuming during inference since they solve convex optimization programs. We use an interior-point solver in this work, which is guaranteed to converge within polynomial

time, i.e. the time to obtain a solution up to precision ε is no more than some polynomial function of ε, N, and M. However, the exact complexity is unknown, which we should expect to be worse than $\Omega((NM)^3)$ (Renegar 1988) (as $\mathcal{O}((NM)^3)$ is the time it would take for the interior-point method to converge if the program were a linear program). We show in Table 1 the training and inference time of the proposed mechanism compared with PF and PA in an actual ten-agent three-resource simulation. While requiring computation in the training process, `ExS-Net` produces allocations much faster during inference.

A mixture of PA and PF provides a stronger benchmark in terms of trading off both NSW and exploitability. Given $\rho \in [0,1]$, we consider the mechanism f^{mixture} below. Varying ρ between $[0,1]$ interpolates between PF and PA in expectation. We set $\rho = 1/2$ for comparison in Section 5.2.

$$r \sim \text{Bern}(\rho),\ f^{\text{mixture}}(v,x,b) = \begin{cases} f^{PF}(v,x,b), & \text{if } r=1 \\ f^{PA}(v,x,b), & \text{if } r=0 \end{cases}$$

Table 1: Training and inference time in a ten-agent three-resource allocation problem (normalized).

Mechanism	Training Time (Theory)	Inference Time (Theory)	Training Time (Simulation)	Inference Time (Simulation)
Proportional Fairness	0	Worse than $\Omega((MN)^3)$	0	40.1
Partial Allocation	0	Worse than $\Omega((MN)^3)$	0	310.0
`ExS-Net` (Proposed)	$\mathcal{O}(MN)$	$\mathcal{O}(MN)$	**1**	**1**

Evaluation Metrics. Mechanisms are evaluated on NSW (2), exploitability, and efficiency

$$\textbf{expl}(f,v,x,b) = (1/N)\sum_{i=1}^{N} \textbf{expl}_i(f,v,x,b), \quad \textbf{efficiency}(f,v,x,b) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{N} f_{i,m}(v,x,b)}{\sum_{m=1}^{M} b_m}.$$

It is desirable for mechanisms to have a high efficiency to reduce the waste of resource, although efficiency merely is a byproduct of NSW in the `ExS-Net` training objective. Since budget may exceed total demand, maximum efficiency can be smaller than 1. The PF mechanism is fully efficient (i.e. all available resources are used for sufficient demand), and hence serves as a suitable reference point.

5.2 Scaling System Parameters

Figure 4 shows the mechanism performance on two-agent two-resource and ten-agent three-resource allocation problems, with the tolerance on exploitability set with respect to that of the PF mechanism to 10^{-3} and 10^{-4}, respectively. The 2x2 system is the smallest non-trivial case, while the 10x3 system is the largest considered in the recent works (Dütting et al. 2024; Ivanov et al. 2022) on auction design with payment.

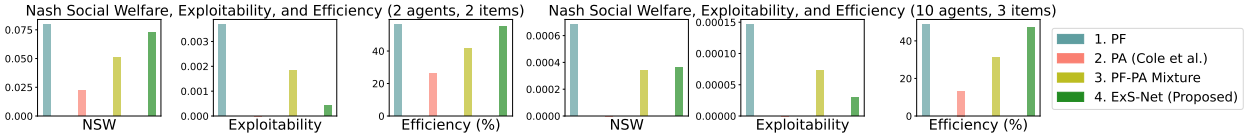

Figure 4: Mechanism performance in 2x2 and 10x3 systems.

The figure shows that `ExS-Net` achieves an advantageous trade-off between PF and PA: it consistently reduces the exploitability of PF by over at least 80% while achieving remarkable NSW and almost full efficiency. Compared with PA, `ExS-Net` significantly improves the efficiency and NSW. In addition, `ExS-Net` outperforms the interpolated mixture of PF and PA under all three metrics.

In Figure 5, we conduct experiments in larger systems with 40-60 agents and plot metrics including the averaged per-agent utility, exploitability, and efficiency. The improvement is consistent – each agent on average gets over 82% of the utility it would get under the PF mechanism while exploitability is reduced by 88-96%.

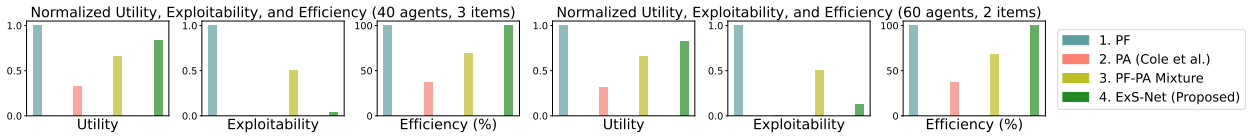

Figure 5: Mechanism performance in 40x3 and 60x2 systems.

5.3 Distribution Mismatch

We investigate performance when the distribution of values and demands the mechanism is trained differs from that on which it is tested. This mismatch may occur due to measurement error or more pernicious sources such as strategically misrepresenting preferences. While we may limit such behavior by running an IC mechanism (such as the PA mechanism) to collect training samples, in this section we numerically examine the robustness of the mechanisms from Section 3 if untruthful training samples are present.

Recall that F denotes the true joint distribution and F' denotes the distribution of the training samples. Further denote by $\omega^\star(F')$ the optimal solution to the ERM problem (5) under samples drawn from distribution F'. With a large discrepancy between F and F, the performance of $\omega^\star(F')$ under the true distribution F, measured by $\mathbb{E}_F[\mathbf{NSW}(\omega^\star(F'),v,x,b)]$ and $\mathbb{E}_F[\mathbf{expl}(\omega^\star(F'),v,x,b)]$, can theoretically be worse than that of $\omega^\star(F)$. Nonetheless, empirically we observe robustness: $\mathbb{E}_F[\mathbf{NSW}(\omega^\star(F'),v,x,b)]$ and $\mathbb{E}_F[\mathbf{expl}(\omega^\star(F'),v,x,b)]$ closely match $\mathbb{E}_F[\mathbf{NSW}(\omega^\star(F),v,x,b)]$ and $\mathbb{E}_F[\mathbf{expl}(\omega^\star(F),v,x,b)]$ across two common types of distribution mismatch.

Randomly perturbed training samples In a 2×2 system we suppose that the distribution F follows (6) while the training samples $\{(v^l,x^l)\}_l$ are generated under the perturbation of a heavy-tailed Cauchy random variable, which is often used to model unobserved strategic behavior.

Adversarially generated training samples Again, we consider a 2×2 setting where the true distribution F is described in (6). Suppose that the training dataset is composed of historical valuations and demands collected through past interactions with the agents. If the agents believe that allocations are made by the PF mechanism in these past interactions, they may have reported strategically with the aim of maximizing their own utility, which means that the training samples $\{(v^l,x^l)\}_l$ will take the following form:

$$\vec{v}^l_{i,m}\sim \text{Unif}(0.1,1),\quad \vec{x}^l_{i,m}\sim \text{Unif}(0.1,1)\quad v^l_i,x^l_i=\arg\max_{v'_i,x'_i} u_i(f((v'_i,\vec{v}^l_{-i}),(x'_i,\vec{x}^l_{-i}),b),\vec{v}^l,\vec{x}^l). \tag{7}$$

Table 2 shows the performance of the proposed mechanisms trained using (7), which closely tracks that under truthful training data across all evaluation metrics. This experiment demonstrates the robustness of the learning-based mechanisms to mismatch between training and inference distribution.

5.4 Trade-Off Frontier between Nash Social Welfare and Exploitability

Instead of training the mechanisms on the constrained objective (4), we can alternatively consider the unconstrained program below, which allows us to directly control the weight of exploitability relative to NSW.

$$\max_{\omega}\ \mathbb{E}_{(v,x,b)\sim F}[\mathbf{logNSW}(f^\omega,v,x,b)+\alpha\textstyle\sum_{i=1}^N \mathbf{expl}_i(f^\omega,v,x,b)]. \tag{8}$$

In Figure 2, we plot the performance frontier of the `ExS-Net` mechanism under as α varies across orders of magnitude in a two-agent two-resource system. The figure shows that the probabilistic mixture of PF and PA, indicated by the dotted line, falls inside the frontier by a large margin, indicating that the proposed `ExS-Net` significantly outperforms the mixture of PA and PF mechanisms in terms of the trade-off between exploitability and NSW.

Table 2: Performance of mechanism trained under data containing untruthful reporting of agents' preferences.

Mechanism	NSW	Exploitability	Efficiency (%)
Proportional Fairness Mechanism	8.00e-2±1.5e-2	3.70e-3±1.2e-3	56.6±3.9
Partial Allocation Mechanism	2.22e-2±2.1e-3	0±0	26.2±2.5
PF-PA Mixture	5.11e-2±6.6e-3	1.85e-3±6.2e-4	41.4±3.6
`ExS-Net` (Trained on Truthful Data)	7.24e-2±4.3e-3	4.04e-4±9.6e-5	55.0±4.2
`ExS-Net` (Trained on Randomly Perturbed Data)	7.26e-2±1.0e-2	4.78e-4±2.9e-4	55.2±1.4
`ExS-Net` (Trained on Adversarial Data)	7.01e-2±2.0e-2	5.25e-4±1.7e-4	54.9±2.6

6 CONCLUSION

We studied learning of fair resource allocation mechanisms without payments. We designed a modular neural network architecture that can take in general layouts (feed-forward, transformers, CNN etc) in the hidden layers and tack-on the strategic resource withholding in the activation layer to achieve the desired trade-off. This flexibility increases the generality and applicability of the proposed methods for approximately fair and incentive compatible mechanisms in the payment-free setting. The payment free setting required changes in the "hardware" and is not a simple flip of the objective in Dütting et al. (2024). Additionally, the welfare in the objective posed additional challenges in characterizing the generalization bound, adding to the literature on learning mechanisms (Dütting et al. 2024; Ivanov et al. 2022; Mishra et al. 2022). We conducted an extensive empirical study of the proposed mechanisms and compared with the state-of-the-art baselines that lie at the end of the spectrum. The results positively affirm that the desired trade-offs are achieved by `ExS-Net`.

DISCLAIMER

This paper was prepared for informational purposes in part by the Artificial Intelligence Research group of JP Morgan Chase & Co and its affiliates ("JP Morgan"), and is not a product of the Research Department of JP Morgan. JP Morgan makes no representation and warranty whatsoever and disclaims all liability, for the completeness, accuracy or reliability of the information contained herein. This document is not intended as investment research or investment advice, or a recommendation, offer or solicitation for the purchase or sale of any security, financial instrument, financial product or service, or to be used in any way for evaluating the merits of participating in any transaction, and shall not constitute a solicitation under any jurisdiction or to any person, if such solicitation under such jurisdiction or to such person would be unlawful.

REFERENCES

Bertsimas, D., V. F. Farias, and N. Trichakis. 2011. "The Price of Fairness". *Operations Research* 59(1):17–31.

Bliem, B., R. Bredereck, and R. Niedermeier. 2016. "Complexity of Efficient and Envy-Free Resource Allocation: Few Agents, Resources, or Utility Levels.". In *International Joint Conference on Artificial Intelligence*, 102–108.

Camacho, F., R. Fonseca-Delgado, R. P. Pérez, and G. Tapia. 2021. "Beyond Identical Utilities: Buyer Utility Functions and Fair Allocations". *arXiv preprint arXiv:2109.08461*.

Cole, R., V. Gkatzelis, and G. Goel. 2013. "Mechanism Design for Fair Division: Allocating Divisible Items without Payments". In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, 251–268.

Dekel, O., F. Fischer, and A. D. Procaccia. 2010. "Incentive Compatible Regression Learning". *Journal of Computer and System Sciences* 76(8):759–777.

Dütting, P., Z. Feng, H. Narasimhan, D. C. Parkes and S. S. Ravindranath. 2024. "Optimal Auctions through Deep Learning: Advances in Differentiable Economics". *Journal of the ACM* 71(1):1–53.

Giannakopoulos, Y. and E. Koutsoupias. 2014. "Duality and Optimality of Auctions for Uniform Distributions". In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, 259–276.

Goktas, D. and A. Greenwald. 2022. "Exploitability Minimization in Games and Beyond". In *Advances in Neural Information Processing Systems*, Volume 35, 4857–4873.

Hartline, J. D. and T. Roughgarden. 2008. "Optimal Mechanism Design and Money Burning". In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, 75–84.

Ivanov, D., I. Safiulin, I. Filippov, and K. Balabaeva. 2022. "Optimal-er Auctions through Attention". In *Advances in Neural Information Processing Systems*, Volume 35, 34734–34747.

Kelly, F. 1997. "Charging and Rate Control for Elastic Traffic". *European Transactions on Telecommunications* 8(1):33–37.

Mishra, S., M. Padala, and S. Gujar. 2022. "EEF1-NN: Efficient and EF1 Allocations through Neural Networks". In *Pacific Rim International Conference on Artificial Intelligence*, 388–401. Springer.

Nisan, N. 2007. "Introduction to Mechanism Design (for Computer Scientists)". *Algorithmic Game Theory* 9:209–242.

Pavlov, G. 2011. "Optimal Mechanism for Selling Two Goods". *The BE Journal of Theoretical Economics* 11(1):1–35.

Procaccia, A. D. and M. Tennenholtz. 2013. "Approximate Mechanism Design without Money". *ACM Transactions on Economics and Computation (TEAC)* 1(4):1–26.

Renegar, J. 1988. "A Polynomial-Time Algorithm, Based on Newton's Method, for Linear Programming". *Mathematical Programming* 40(1-3):59–93.

AUTHOR BIOGRAPHIES

SIHAN ZENG is a Research Scientist at JPMorgan AI Research with expertise in optimization and reinforcement learning. His email address is sihan.zeng@jpmchase.com.

SUJAY BHATT is a Research Lead at JPMorgan AI Research with expertise in bandits, reinforcement learning, and optimization. His email address is sujay.bhatt@jpmchase.com.

ELEONORA KREACIC is a Research Lead at JPMorgan AI Research, with interest and expertise in generative models and graph theory. Her email address is eleonora.kreacic@jpmorgan.com.

PARISA HASSANZADEH is a AI Research Scientist at Samsung SDS. Her research broadly spans telecommunications, machine learning, and the applications of AI in finance. Her email address is parisah@nyu.edu.

ALEC KOPPEL is a Research Lead at JPMorgan AI Research with expertise in optimization, machine learning, and sequential decision making. His email address is alec.koppel@jpmchase.com.

SUMITRA GANESH is a Research Director at JPMorgan AI Research, leading a team on AI Agents, multi-agent simulations, and sequential decision making. Her email address is sumitra.ganesh@jpmorgan.com.