TECHNISCHE UNIVERSITÄT MÜNCHEN

TUM School of Management

Essays on Bidding Strategies and Auction Design

Paul R. Sutterer

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Volkswirtschaftslehre (Dr. oec. publ.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Sebastian Schwenen
Prüfer der Dissertation:	1. Prof. Dr. Martin Bichler
	2. Prof. Dr. Rainer Kolisch

Die Dissertation wurde am 16.12.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15.04.2021 angenommen.

Abstract

Auctions have become a popular means for the distribution of goods when an information asymmetry between buyers and sellers makes setting a fixed price difficult. Two central topics in auctions are optimal bidding strategies and auction design. They address the questions of how high a bidder should bid and how an auction should be designed to incentivise certain bidder behavior, e.g. truthful bidding. In this dissertation, we address these questions with three contributions.

Online advertising represents one of the biggest advertising markets worldwide. In display ad auctions, bidders bid on the opportunity to show their ad to a user. These auctions are performed as second-price auctions which are known to be incentive-compatible, i.e. truthful bidding is a weakly dominant strategy. However, advertisers are typically limited by budgets or campaign targets, as a result preferences are no longer separable and the auction is not incentive-compatible. We study the effects of truthful bidding on efficiency and bidders' utility in these auctions. We find that efficiency can be as low as 50% in the worst-case, while in fact it is close to 100% on average. Bidders can gain by deviating from truthful bidding only if there is little competition. Otherwise, truthful bidding presents a good strategy.

Another large market, and one of this century's major challenges, is the expansion of renewable energy sources. Governments worldwide use auctions to tender renewable energy capacities in exchange for guaranteed remuneration prices. In Germany, nation wide auctions are used for the expansion of onshore wind power plants since 2017. Lacking to set the right incentives, they have led to huge discrepancies in the distribution of capacities between the north and the south. Missing grid capacity causes bottlenecks, the need for redispatch, and inefficiencies. Furthermore, larger project developers cannot communicate potential synergies and each individual project receives an individual price that can be perceived as unfair and intransparent. We develop and evaluate an alternative combinatorial auction design that allows to define regional capacities and allows bidders to communicate their synergies via bundle bids. Considering the size of the market, we compute linear and approximate anonymous Walrasian prices per region which are strategy-proof in the large. Due to the indivisibility of items Walrasian prices might not always exist. In this case we introduce minimal personal markups. In a counterfactual analysis we show that the alternative auction design implements an efficient allocation while hardly increasing remuneration prices and therefore significantly decreasing the overall cost for the tax payer. The presented auction design is policy relevant and can be effectively implemented.

In smaller markets with fewer participants, strategic interaction between bidders becomes more relevant. In fact, many auctions are not incentive-compatible and bidders have to decide their best bidding strategy. Bayes-Nash equilibrium presents a central solution concept to such Bayesian games. However, Bayes-Nash equilibria are known for only very few and simple auctions and finding them is difficult, in fact at least PPAD-complete. We develop a learning algorithm based on neural networks that we call Neural Pseudogradient Ascent (NPGA). NPGA learns Bayes-Nash equilibria based on iterative self-play and evolutionary strategies. We show that NPGA is able to learn approximate Bayes-Nash equilibria in a wide variety of auction settings, including single-item, multi-item, and combinatorial auctions. In auctions for which no Bayes-Nash equilibrium is known, NPGA may provide an estimate for the ε-Bayes-Nash equilibrium by computing best responses. It demonstrates to be a powerful tool for economists and practitioners to solve auction games.

Zusammenfassung

Auktionen stellen ein zunehmend beliebtes Instrument zur Verteilung von Gütern dar, insbesondere wenn Informationsasymmetrien zwischen Käufern und Verkäufern die Bestimmung fixer Preise erschweren. Zwei zentrale Problemstellungen im Zusammenhang mit Auktionen sind das Auktionsdesign und die optimale Bietstrategie. Konkret stellen sich die zwei Fragen: (1) wie hoch jeder Bieter mit seinen Geboten gehen sollten und (2) wie eine Auktion gestaltet sein sollte, um ein bestimmtes Bieterverhalten zu fördern, wie beispielsweise wahrheitsgemäßes Bieten. In dieser Dissertation behandeln wir beide Fragen in drei Beiträgen.

Onlinewerbung steht inzwischen für einen Großteil des weltweiten Werbemarktes. In *Display-Ad-Auktionen* bieten Werbetreibende darauf, ihre Werbung einem Nutzer zeigen zu können. Diese Auktionen werden als Zweitpreisauktionen durchgeführt, welche bekanntermaßen anreizkompatibel sind, womit wahrheitsgemäßes Bieten eine schwach dominante Strategie darstellt. Allerdings sind Werbetreibende oft Budgetbeschränkungen oder Kampagnenzielen unterworfen. Daraus ergibt sich, dass Präferenzen nicht mehr unabhängig voneinander und die Auktionen nicht länger anreizkompatibel sind. Wir untersuchen die Effekte von wahrheitsgemäßem Bieten im Hinblick auf Effizienz und den Bieternutzen in diesen Auktionen. Wir zeigen, dass die Effizienz im schlechtesten Fall bis auf 50% sinken kann, im Durchschnitt allerdings bei nahezu 100% liegt. Lediglich bei geringem Wettbewerb können Bieter durch das Abweichen von wahrheitsgemäßem Bieten profitieren. Andernfalls zeigt sich, dass wahrheitsgemäßes Bieten eine gute Strategie ist.

Einen weiteren großen Markt, und eine der großen Aufgaben dieses Jahrhunderts, stellt der Ausbau von erneuerbaren Energien dar. Weltweit nutzen Regierungen Auktionen für den Kapazitätsausbau von erneuerbaren Energien, in welchen sie Projektentwicklern garantierte Vergütungspreise anbieten. In Deutschland werden seit 2017 Auktionen auf nationaler Ebene für den Ausbau der Windenergie zu Land genutzt. Ein mangelhaftes System der Anreizsetzung hat in der Vergangenheit allerdings zu einer ungleichen Verteilung der Kapazitäten zwischen Nord und Süd geführt. Die mangelnde Netzkapazität führt zu Engpässen, notwendigen *Redispatch*-Maßnahmen und Ineffizienzen. Darüber hinaus lassen sich potenzielle Synergieeffekte großer Projektentwickler im derzeitigen Auktionsdesign nicht abbilden und jedes Projekt erhält jeweils einen individuellen Vergütungspreis. Dies kann als unfair und intransparent wahrgenommen werden. Wir entwickeln und evaluieren ein alternatives, kombinatorisches, Auktionsdesign. Dieses erlaubt es dem Auktionator, regionale Kapazitäten zu bestimmen, als auch den Bietern ihre Synergien durch Bündelgebote abzubilden. Angesichts der Größe des Marktes berechnen wir für jede Region lineare und annähernd anonyme *Walrasian prices*, welche *strategy-proof in the large* sind. Aufgrund der Unteilbarkeit der Güter kann es allerdings sein, dass die Bestimmung eines solchen Preises nicht möglich ist. In diesem Fall verwenden wir minimale persönliche Aufschläge. Eine kontrafaktische

Analyse zeigt, dass das alternative Auktionsdesign zu einer effizienten Allokation führt und gleichzeitig kaum die Vergütungspreise erhöht. Entsprechend führt das Design zu einer signifikanten Reduktion der Gesamtkosten für den Steuerzahler. Das vorgestellte Auktionsdesign ist politisch relevant und faktisch umsetzbar.

In kleineren Märkten mit weniger Marktteilnehmern ist strategisches Verhalten der Bieter von größerer Bedeutung. Tatsächlich sind viele Auktionen nicht anreizkompatibel und entsprechend müssen Bieter sich die Frage stellen was ihre beste Bietstrategie ist. Ein zentrales Lösungskonzept solcher bayesianischen Spiele liegt in der Berechnung eines Bayes-Nash Gleichgewichts. Dieses Gleichgewicht ist allerdings nur für wenige, sehr einfache Auktionen bekannt, und es zu finden ist schwer – mindestens PPAD-vollständig. Basierend auf neuronalen Netzen entwickeln wir einen Lernalgorithmus den wir *Neural Pseudogradient Ascent* (NPGA) nennen. NPGA lernt Bayes-Nash Gleichgewichte durch iteratives Spielen gegen sich selbst sowie durch die Berechnung von Gradienten mithilfe von *evolutionary strategies*. Wir zeigen, dass NPGA in der Lage ist, in einer Reihe von Auktionsumgebungen, einschließlich *single-item*, *multi-item* und *combinatory* Auktionen, approximative Bayes-Nash Gleichgewichte zu erlernen. Für Auktionen ohne bekanntes Bayes-Nash Gleichgewicht ermittelt NPGA einen Schätzer für ε-Bayes-Nash Gleichgewichte, indem er *best response* Strategien berechnet. NPGA präsentiert sich sowohl für Ökonomen als auch für alle Beteiligten einer Auktion als bedeutendes Hilfsmittel zur Lösung von Auktionsspielen.

Acknowledgements

This thesis would not have been possible if it was not for a number of great people who not only supported me during the time of my doctorate, but that also helped pave the way for me to pursue it.

First and foremost, I am grateful to my supervisor Martin Bichler for his constant, unlimited support and great collaboration over the last years. He always found the time to provide good advice and perspective.

A great thanks go to Rainer Kolisch for not only reviewing my thesis, but all the more so for supporting me during most of my Bachelors and Masters studies, as well as my doctorate. It is fair to say that he played a major role in setting the path to my academic career.

I am grateful to my colleagues that were never shy of offering support. I would especially like to thank my mentor Stefan Waldherr and my colleague Stefan Heidekrüger not only for examining my thesis and providing constructive feedback but for great collaboration over the years, always with a good laugh, too.

Last but not least, I want to thank my family and my wife for constant support over the years, especially during the last one. I am very grateful for their never ending support and guidance.

Contents

1 Introduction

For over two thousand years, auctions have been used for the distribution of goods. Technology advancements during the last decades, leading to the rise of electronic markets, enabled a rapid increase in the application of auctions. In the course of these developments, auctions have gained increasing interest by researchers as well as practitioners. Some of the most prominent areas of application today are online advertisement markets (Chen and Stallaert, 2014; De et al., 2010; Yang and Ghose, 2010; Zhao and Xue, 2012), spectrum sales (Bichler and Goeree, 2017) and the expansion of renewable energy sources (Winkler et al., 2018). The contemporary importance of auctions has most recently been highlighted by Paul Milgrom and Robert Wilson being awarded with the Nobel Prize in Economic Sciences for their contributions to auction theory.

Auction markets inherently involve incomplete information about competitors and strategic behavior of market participants. Understanding decision making in such markets is imperative for the efficient, i.e. welfare maximizing, allocation of goods. This thesis' contribution is threefold, presented in four essays that study bidding strategies and auction design in auction markets.

In the first essay, we study the online advertisement market, in particular the display ad auctions (Sutterer et al. (2019), Chapter 3). In this market millions of auctions are performed each day. We find that truthful bidding in such a large market is a good strategy when bidders encounter medium or high competition. We continue with a second essay in large auction markets and study the renewable energy market (Bichler et al. (2020), Chapter 4), in particular auctions for the expansion of onshore wind power plants. Each year, about one hundred bidders participate in these auctions offering to build more than one hundred wind power plants. We propose an alternative auction design that is more efficient and provides linear, near anonymous prices, and is therefore more transparent than the currently applied design. It hardly increases remuneration prices but reduces total costs for the tax payer.

In many auction markets, including display ad auctions with little competition (Sutterer et al., 2019), truthful bidding is not a good strategy. In essays 3 and 4, we develop a learning algorithm for general auction markets. These two essays are consecutive peer reviewed workshop publications focusing on normal-form games and single-item auctions (Heidekrüger et al. (2019), Chapter 5), and on more complex combinatorial auctions (Heidekrüger et al. (2020a), Chapter 6). This research has also been accepted and presented at other peer-reviewed workshops (Heidekrüger et al., 2020b) and a combined journal version of the individual workshop contributions is currently under review. The developed learning algorithm is able to compute Bayes-Nash equilibria for many of these auction markets helping bidders to find a good strategy. We now introduce each of the three topics in the following sections.

1.1 Truthful Bidding in Display Ad Auctions

In 2019 online advertising markets represented about half of the total media spending world wide (emarketer, 2019), effectively making it the most important market for advertising. Next to keyword auctions, display ad auctions have become the predominant means to allocate advertisements to users who visit websites. The opportunity to show an advertisement to such a user is referred to as an impression. Every time a user visits a website, an auction for such an impression is triggered, leading to many millions of auctions every day, each conducted within milliseconds. Typically these auctions are performed as second price auctions, meaning that the winner pays a price in the amount of the second highest bid. Given a bidder's quasi-linear utility function and separability of valuation for items, i.e. the valuation for an item being independent of any previous allocations, this payment rule is known to be incentive-compatible. As a consequence, truthful bidding is a weakly dominant strategy for bidders.

However, advertisers usually have decreasing marginal valuation for additional impressions to win, i.e. the valuation for the first impression is not the same as the valuation after winning thousands of impressions. We consider an economic model where bidders have quasi-linear utility functions but are restricted by campaign targets. This models such a valuation function with decreasing marginal valuation, making a sequence of second-price auctions no longer incentive-compatible. The question is how high a bidder's payoff and the auction's efficiency is when bidders continue to bid truthfully, and how much they can gain by deviating from truthful bidding.

In this study, we show that the efficiency can theoretically be as low as 50%. However, numerical experiments reveal that this estimate is far too pessimistic. The average efficiency is close to 100% throughout all experiments. Furthermore, the bidders' payoff is relatively high and by deviating from truthful bidding, bidders risk being worse off when they encounter medium or high competition. Only with few competing bidders is their payoff relatively low and they can gain by lowering their bid. In general, truthful bidding can therefore be a good strategy.

1.2 An Alternative Auction Design for Renewable Energy Auctions

In the second essay, we study another large market, namely the renewable energy market. The expansion of renewable energy sources (RES) has gained increasing importance and a welfare maximizing allocation is crucial for its success. In Europe and Worldwide, auctions are used to determine the remuneration for RES (Winkler et al., 2018), i.e. an amount of money per kilowatt hour (kWh) that auction winners receive for building a renewable energy plant. Bidders typically build renewable energy plants at very productive sites, potentially far-off the main load centres (Grimm et al., 2017). This leads to an inefficient allocation because the renewable energy plants' feeding into the grid must be decreased when supply and demand are not balanced and transmission line capacities are restricted, however, not considered in the allocation. This long term inefficient system configuration is subsidized by the tax payer. They have to bear the costs for additional grid expansion or for the electricity that is produced but not fed into the grid. Moreover, in the current auction design, larger project developers can only submit bids on individual renewable energy plants, i.e. projects, and receive individual remuneration prices.

For the example of the German onshore wind energy auctions, we propose an alternative auction design that accounts for economies of scale for larger project developers, by allowing bundle bids, and that leads to an efficient allocation of wind power plants by implementing regional target capacities (Grimm et al., 2017). The allocation is determined by solving a binary linear program that selects the most cost efficient projects bundles while keeping the exceeding capacity in each region to a minimum.

The proposed combinatorial auction provides a simple pricing rule close to Walrasian prices and maintains a high level of competition between bidders by permitting package bids. Walrasian prices are strategy-proof in the large (Azevedo and Budish, 2019), i.e. approximately incentive-compatible.

In numerical experiments, we evaluate the combinatorial auction compared to three other RES auction designs: the current German nationwide auction design, a simple nationwide auction, and regional auctions. We find that the combinatorial auction design implements a fully efficient allocation due to the possibility of precisely steering the tendered capacities. At the same time, leveraging the possibility to directly communicate synergy effects, project developers can offer bundles of wind power plants for lower prices. The resulting remuneration prices remain low and hardly increase compared to the current auction design. Considering the overall cost for the tax payer, the combinatorial auction design can decrease these costs significantly.

1.3 Computing Bayes-Nash Equilibrium Strategies in Auctions

This thesis' first two essays consider large markets for which truthful bidding often is a good strategy or the mechanism is even strategy-proof in the large. However, in many smaller auctions, or even in the large display ad auctions with little competition (Sutterer et al., 2019), bidders that bid truthfully do overpay.

Bayes-Nash equilibrium is a central solution concept for Bayesian Games such as auctions. Unfortunately, they are known only for some single-item auctions and for very few multi-item auctions. Computing Bayes-Nash equilibria is at least PPAD-complete, as computing Nash-equilibria is already PPAD-complete (Daskalakis et al., 2009). Being able to compute Bayes-Nash equilibria for more complex auction designs could significantly improve the welfare when performing these auctions.

In essays 3 and 4, we consecutively develop a learning algorithm first called Neural Self-Play, later referred to as Neural Pseudogradient Ascent. Neural Pseudogradient Ascent is an iterative learning algorithm based on neural networks to compute Bayes-Nash equilibria in general auction markets via self-play. Each bidders' strategy is modeled by a neural network, using a bidder's private information as input and computing a bidder's bids as output. Leveraging modern GPU hardware enables us to compute batches of the same auction but with many different valuations for bidders all at once. In each learning iteration, we draw batch size many valuations for each bidder, compute their corresponding bids, evaluate their payoffs, and update the networks' parameters based on this feedback.

To update the parameter vector, we cannot use standard backpropagation. Exact gradients of the ex-post utility for a fixed valuation and opponent strategy profiles lead to problems in gradient updates. The bidder's ex-post utility will generally be discontinuous in her action as allocations are discrete. Therefore, we use evolutionary strategies (Salimans et al., 2017) and create random perturbations of the parameter vector by adding Gaussian noise terms. The resulting perturbed models are evaluated with respect to their performance "fitness". The model is ultimately updated in the direction of the weighted average of the noise vectors, with more desirable perturbations being weighted higher than less desirable ones.

First, we apply Neural Pseudogradient Ascent to a number of normal form games to compare its performance to classical learning algorithms in these games with discrete state and action space, namely Fictitious Play and Smooth Fictitious Play. After, we apply Neural Pseudogradient Ascent to the main Bayesian settings of different single-item and combinatorial auctions with various value models and payment rules. For some of these settings Bayes-Nash equilibria are known while for others they are not.

In numerical experiments, we confirm that Neural Pseudogradient Ascent performs similar to Fictitious Play and Smooth Fictitious Play when computing Nash equilibria in normal-form games. In auction games it is able to compute approximate Bayes-Nash equilibria for each of the settings we considered.

1.4 Outline

The remainder of this thesis is structured as follows: In the next chapter (Chapter 2), we introduce the theoretical basis for the methodologies applied within this thesis. We begin by introducing complete information normal-form games and continue with auctions as games of incomplete information with continuous state and action space. We end with introducing neural networks and evolutionary strategies.

In Chapter 3, we present the first essay on truthful bidding in online advertisement markets, followed by the second essay on market design in renewable energy auctions in Chapter 4. In Chapter 5 and 6, we present two consecutive essays in which we present Neural Pseudogradient Ascent as a learning algorithm to learn Nash-equilibria in normal-form games and Bayesian-Nash equilibria in single-item and combinatorial auctions.

Finally, we discuss this thesis' overall contribution to the literature of auction theory in Chapter 7 and conclude in Chapter 8.

2 Theoretical Background

In this chapter, we illustrate the theoretical concepts this thesis is built on and that are applied herein. We begin with complete information games and some of their solution concepts, i.e. Nash equilibria and a learning algorithm called (Smooth) Fictitious Play (Brown, 1951), originating from the economics branch of research. We continue by introducing incomplete information games, i.e. Bayesian games, and the extension of Nash equilibrium to Bayes-Nash equilibrium. Afterwards, we introduce auctions as Bayesian games and describe several auction payment rules and their properties applied in this thesis. Finally, we introduce neural networks, much studied in the computer scientists' branch of research, as a concept to enable learning in Bayesian games.

The fundamental theory in this chapter is largely based on Bichler (2017), Krishna (2009), Fudenberg and Levine (1999), Fudenberg and Levine (2009) and Goodfellow et al. (2016). We lay out the theory and describe how our work build upon it.

2.1 Fundamentals of Game Theory

In this section, we introduce the basics of non-cooperative game theory, beginning with games of complete information. This mostly builds on Bichler (2017), Fudenberg and Levine (1999) and Fudenberg and Levine (2009).

Non-cooperative game theory studies the strategic interaction between self-interested rational agents, in the following referred to as *players*, maximizing their outcome. In this context, strategic means that each player makes a decision about her action aware that it affects the other players, in the following referred to as *opponents*. Interaction describes the interdependence of actions, i.e. each player's outcome depends on her own actions as well as the actions of her opponents; self-interested means each player is only interested in her own outcome; and rational means players are free of emotional biases and pure utility maximizers instead. One of the most well known examples of a complete information game is the prisoner's dilemma:

Example 2.1: Prisoner's Dilemma. *Two suspects are accused of jointly having committed a major crime and are being separately and simultaneously interrogated. Both have a decision to make either confessing (C) to the accusations or denying (D) them. The suspects, in the following the players $\mathcal{I} = \{1, 2\}$, have the same set of actions available $\mathcal{A}_1 = \mathcal{A}_2 = \{C, D\}$. If both suspects confess, they receive a sentence of 4 years' detention each; if both deny, they still receive a sentence of 2 years' detention each for lesser crimes they committed. If one suspect chooses to confess and the other to deny, the former acts as the principal witness and is sentenced to only 1 year, while the latter is sentenced to 5 years.*

These games of complete information in which players need to choose their actions simultaneously are called normal-form games and can best be represented in matrix form (see Table 2.1), where player 1 is the row player and player 2 the column player. The first entry represents the (negative) payoff for player 1 and the second entry the (negative) payoff for player 2.

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

Table 2.1: Payoff Matrix in Prisoners Dilemma

Definition 2.1: Normal-Form Game. *A finite normal-form game with n players can be described as a tuple ($\mathcal{I}$, $\mathcal{A}$, u).*

- *$\mathcal{I}$ is a finite set of n players indexed by i.*
- *$\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$, where $\mathcal{A}_i$ is a finite set of actions available to player i. A vector $b = (b_1, \ldots, b_n) \in \mathcal{A}$ is referred to as an action profile.*
- *$u = (u_1, \cdots, u_n)$, where $u_i : \mathcal{A} \mapsto \mathbb{R}$ is a payoff or utility function for player i.*

As all players are a payoff maximizers they try to find a strategy that maximizes their own utility u_i. These strategies can either represent a particular action, i.e. a *pure strategy*, or a probability distribution over a set of actions, i.e. a *mixed strategy*.

Definition 2.2: Mixed Strategy. *In a normal-form game $(\mathcal{I}, \mathcal{A}, u)$ the set of mixed strategies for player i is $\Sigma_i = \Delta(\mathcal{A}_i)$, where $\Delta(\mathcal{A}_i)$ is the set of all probability distributions (aka lotteries) over $\mathcal{A}_i$.*

The probability that an action b_i is played in strategy β_i is denoted as $\beta_i(b_i)$. When players choose to play a mixed-strategy profile $(\beta_1, \cdots, \beta_n)$, the expected utility of player i can be described as $u_i(\beta) = \sum_{b \in \mathcal{A}} u_i(b) \prod_{j=1}^{n} (\beta_j(b_j))$. Notice that the term in parentheses, $\prod_{j=1}^{n} (\beta_j(b_j))$, is the probability that action profile $b = (b_1, \cdots, b_n)$ will be played. A mixed-strategy profile is the Cartesian product of the individual mixed-strategy sets, $\Sigma_1 \times \cdots \times \Sigma_n$ (Bichler (2017), p. 12). Players can either play pure strategies or mixed strategies, while in general any pure strategy can also be represented as a mixed strategy with $\beta_i(b_i) = 1$ for a certain action b_i.

In the next section, we introduce solution concepts for normal-form games that help players play optimal strategies and can serve as predictions for the outcome of a game with rational players.

2.1.1 Solution Concepts: Equilibria

When players in a normal-form game have to choose a strategy, it is not always straightforward what the best strategy is. In economics, several solution concepts exist for these games that are often referred to as equilibria. An equilibrium denotes a stable state as a result of a dynamic process in which the players' behavior is consistent and they have no incentive to change their behavior (Dixon (2001), p. 24). Therefore, each player is maximizing their own utility by remaining in this state.

While a few notions of equilibria exist, for this thesis we focus on four that are most relevant in our studies, namely: dominant-strategy equilibrium, Nash equilibrium, Bayes-Nash equilibrium and Walrasian equilibrium (prices).

In a dominant-strategy equilibrium, a player's strategy is optimal regardless of what the opponents' strategies are. Formally:

Definition 2.3: (Weakly) Dominant Strategy. *A strategy β_i is called a (weakly) dominant strategy if it is always better (at least equally good) than any other strategy $\beta_i' \in \Sigma_i$, regardless any of the possible opponents' strategies.*

Definition 2.4: Dominant-Strategy Equilibrium. *If each player chooses to play their dominant strategy, the corresponding strategy profile β denotes a dominant-strategy equilibrium.*

While in a dominant-strategy equilibrium the opponents' strategies β_{-i} have no impact on the decision of player i for strategy β_i, they can still affect player i's payoff.

Dominant-strategy equilibria are rare since a player's optimal strategy often depends on the opponents' strategies. A more frequent, and one of the most significant solution concepts, is the Nash equilibrium (Nash et al., 1950). A Nash equilibrium exists when no player can increase their payoff by deviating from their current strategy given the current strategies of all opponents, i.e. each player plays their best-response; or formally:

Definition 2.5: Nash Equilibrium. *A strategy profile $\beta = (\beta_1, ...\beta_n)$ is a Nash equilibrium if for each $i \in \mathcal{I}$: the strategy β_i is a best response to the opponents strategies β_{-i}; therefore, the utility $u_i(\beta_i, \beta_{-i}) \geq u_i(\beta'_i, \beta_{-i})$, $\forall \beta'_i \in \Sigma_i$.*

Both dominant-strategy equilibrium and Nash equilibrium can exist as *pure* as well as *mixed* equilibria. Each dominant-strategy equilibrium is also a Nash equilibrium, however the reverse is not necessarily true. Note that neither of them is guaranteed to be Pareto optimal, meaning that all players could still be better off by playing a different strategy profile β' all together. If in β' again no player would have an incentive to change their strategy then β' denotes another equilibrium. Otherwise, it is not an equilibrium and players would not remain in β'.

Definition 2.6: Pareto Optimality. *Any strategy profile that is not Pareto-dominated is Pareto optimal. A strategy profile β Pareto-dominates strategy profiles β' if, for all $i \in \mathcal{I}$, $u_i(\beta) \geq u_i(\beta')$ and there exists a $j \in \mathcal{I}$ for which $u_j(\beta) > u_j(\beta')$.*

In Example 2.1, for both players to *confess* is a dominant-strategy equilibrium (therefore also a Nash equilibrium) since it is always better to *confess*, regardless the opponent's action. However, the Pareto optimal solution would be for both players to *deny*. They would both serve 2 instead of 4 years. Yet, to *deny* is not stable because each player would be better off by individually deviating and so it is not an equilibrium.

Finding dominant-strategy equilibria is not as hard as finding Nash equilibria because of the dependencies on opponents' strategies. In fact, finding a Nash equilibrium has shown to be PPAD-complete (Daskalakis et al., 2009). In the next section, we introduce two classical learning algorithms to solve normal-form games and potentially to find Nash equilibria.

2.1.2 Learning in Complete Information Games with Discrete State and Action Space

The theory of learning in games explores how, which, and what kind of equilibria might arise as a consequence of an iterative learning process. Fudenberg and Levine (1999) provide deep insights into the theory of learning and their more recent article (Fudenberg and Levine, 2009) complements the book well. In the following, we briefly present two fundamental learning algorithms, first introduced by Brown (1951), namely Fictitious Play and its extension Smooth Fictitious Play.

Fictitious Play

Fictitious Play is an iterative learning algorithm that adapts a player's strategy based on past observations of her opponents' actions. While normal-form games are typically played only once, Fictitious Play needs multiple iterations to be able to learn and potentially converge. Therefore, it has many interpretations, one of which is that it can be considered a fictitious pre-play by each player to learn more about a game's dynamics.

Definition 2.7: Fictitious Play. *Fictitious Play is an iterative learning algorithm. Each player begins with initial beliefs regarding the opponents' strategies: $\hat{\beta}_{-i}$. Given these beliefs, each player chooses an action $b_i \in A_i$ that maximizes her expected utility given her beliefs, i.e. plays a best-response:*

$$b_i = \arg\max_{b \in \mathcal{A}_i} \mathbb{E}[u_i(b, \hat{\beta}_{-i})]$$

In the next iteration, beliefs about the opponents' strategies $\hat{\beta}_{-i}$ are updated by the actual observed actions b_{-i}.

Fictitious Play always determines one action to be played and therefore it can only converge to pure Nash equilibria. If only a mixed Nash equilibrium exists, the actions chosen by Fictitious Play will oscillate. However, the empirical distribution of actions chosen by Fictitious Play can still converge. In this case, the strategy profile β is a mixed Nash equilibrium (Fudenberg and Levine (1999), p. 44). Therefore, usually the empirical distribution of actions is considered in Fictitious Play rather than the actual actions that are played. Fictitious Play does not converge in general (Shapley, 1964) but has been shown to do so in some general settings. For details about convergence guarantees in Fictitious Play, we refer the interested reader to Fudenberg and Levine (1999).

Smooth Fictitious Play

While actions are chosen deterministically in Fictitious Play, Smooth Fictitious Play provides a probability distribution over actions from which one action is randomly drawn. There are multiple motivations for this extension. On the one hand, players are less exploitable due to the inclusion of uncertainty about the final action; on the other hand, a slight change in the observations does not lead to an abrupt change in the strategy but players randomize when they are nearly indifferent between actions.

Definition 2.8: Smooth Fictitious Play. *Smooth Fictitious Play is an extension of Fictitious Play in which player i's actions are drawn according to a best-response probability distribution. One possible representation is with a logistic best-response such that the probability of player i to play action b_i given her beliefs $\hat{\beta}_{-i}$ is:*

$$\beta_i(b_i|\hat{\beta}_{-i}) = \frac{e^{\tau \cdot u_i(b_i, \hat{\beta}_{-i})}}{\sum_{r_i \in \mathcal{A}_i} e^{\tau \cdot u_i(r_i, \hat{\beta}_{-i})}}$$

The parameter τ is sometimes referred to as the temperature and determines the level of stochasticity applied or the level of indifference between actions. When $\tau \to 0$ the player is nearly indifferent between actions, and when $\tau \to \infty$ the player's probability of choosing the best response approaches 1.

As Smooth Fictitious Play considers a probability distribution over actions, it can converge to a mixed Nash equilibrium in actual play, i.e. β, in contrast to Fictitious Play. For details about convergence results in Smooth Fictitious Play we refer the interested reader to Fudenberg and Levine (2009).

In Chapter 5 (Heidekrüger et al., 2019) we implement Fictitious and Smooth Fictitious Play for a number of normal-form games with pure and mixed Nash equilibria and compare the results to our own learning algorithm. Finally, we build on fundamentals in game theory in Chapter 6 (Heidekrüger et al., 2020a).

2.2 Auction Theory

In the previous section, we have introduced the fundamentals of game theory and complete information normal-form games. However, often players are aware of only their own utility function u_i and the set of all players' utility functions u is unknown to them, while a probability distribution for u is common knowledge. This results in games of incomplete information, i.e. Bayesian games. One of the most prominent applications of such games are auctions.

An auction can be considered a mechanism in which usually one, but possibly more, players, referred to as auctioneers, offer a good, that is tangible or intangible, for sale. A number of players, referred to as bidders, compete for this good by submitting bids. Typically, the bidder with the highest bid wins the good. In the following we refer to a good as an item. The winner pays a monetary amount determined by some payment rule.

In the following, we formally introduce auctions as a form of Bayesian games, followed by the most relevant payment rules, applied in this thesis, and their properties. This section builds on Bichler (2017) and Krishna (2009) who provide more detailed information on the topics at hand.

2.2.1 Auctions as Bayesian Games

Throughout this thesis, we restrict ourselves to auctions with one auctioneer who does not participate as a player in the game but solely offers the item for sale without any strategic interaction. We follow the standard independent private value framework and assume bidders' valuations as known only to themselves, drawn from a common probability distribution and independent of each other (only in Chapter 6 we consider other value models). The auctioneer has no valuation for any of her items. For the following definitions, we consider an auction of only a single item:

Definition 2.9: Auctions as Bayesian Games. *An auction can be described as a Bayesian game by a tuple $(\mathcal{I}, \mathcal{A}, \mathcal{V}, u, F)$ with:*

- *$\mathcal{I} = \{1, ..., n\}$ is a set of n players, i.e. bidders;*

- *$\mathcal{A}_i$ is the set of actions, i.e. bids, available to bidder i, with $b_i \in \mathcal{A}_i$ being a specific bid and $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$;*

- *$\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_n$ is the set of all possible types, i.e. valuations that bidders potentially have for an item. $\mathcal{V}_i$ is the valuation space of bidder i and v_i represents the actual valuation that bidder i has for an item;*

- *$u_i : \mathcal{A} \times \mathcal{V}_i \mapsto \mathbb{R}$ or $u_i(b_1, ..., b_n; v_i)$ is the payoff or utility function for player i that depends on the bid profile $b = (b_1, ..., b_n)$ and the valuation of i.*

- *$F : \mathcal{V} \to [0, 1]$ defines a common prior probability distribution of the players' types and is assumed to be common knowledge among all players.*

In addition to complete information games, we now consider player types $\mathcal{V}$, i.e. valuations in the context of auctions, based on a probability distribution over those types F that is common knowledge. Due to this extension, we can have three different perspectives on the game, namely: ex-ante, ex-interim and ex-post. Each perspective represents the game at a different point in time: ex-ante - none of the players knows their own or the opponents' types, i.e. valuations, but only the common prior F; ex-interim - each player knows their own valuation but not the opponents' valuations; ex-post - each player knows the actual outcome of the game.

It is now possible to define a strategy as mapping of a type to an action, or as a valuation to a bid: $\beta_i : \mathcal{V}_i \mapsto \mathcal{A}_i$ and $\beta_i(v_i) = b_i$. The ex-interim utility is now given by the expected utility over the belief of the opponents' types and a player's own type v_i:

$$\mathbb{E}(u_i(\beta_i, \beta_{-i}, v_i)|\beta_{-i}, v_i) = \sum_{v_{-i} \in V_{-i}} u_i(\beta_i(v_i), \beta_{-i}(v_{-i}), v_i)p(v_{-i}|v_i),$$

with $p(v_{-i}|v_i)$ being the conditional probability that the opponents' valuation profiles are v_{-i}.

The notion of a Nash equilibrium now easily extends to that of a Bayes-Nash equilibrium as set out below:

Definition 2.10: Bayes-Nash Equilibrium . *A Bayes-Nash equilibrium is the Nash equilibrium of a Bayesian game, i.e.,*

$$\mathbb{E}(u_i(\beta_i, \beta_{-i}, v_i)|\beta_{-i}, v_i) \geq \mathbb{E}(u_i(\beta_i', \beta_{-i}, v_i)|\beta_{-i}, v_i)$$

for all $\beta_i'(v_i)$ and for all types v_i occurring with positive probability.

Often Bayes-Nash equilibria are not known and finding them is very difficult. Instead sometimes ε-Bayes-Nash equilibria are reported which require the computation of best responses for each player. It is defined as:

Definition 2.11: ε-Bayes-Nash Equilibrium. *A strategy profile β forms an ε-Bayes-Nash equilibrium if no player $i \in I$ can improve her utility more than ε by unilaterally changing her strategy β_i to playing a best-response.*

Note that an ε-Bayes-Nash equilibrium can be arbitrarily far from a Bayes-Nash equilibrium with respect to players' utilities as well as with respect to the corresponding strategies β. A 0-Bayes-Nash equilibrium, however, is always identical to a Bayes-Nash equilibrium.

2.2.2 Efficient Allocation in Auctions

So far we have only considered single items. To generalize the possibility of multiple items being sold at once, we refer to a single item from a set of all items as $k \in \mathcal{K}$ and to $S \subseteq \mathcal{K}$ as a subset or bundle of items from $\mathcal{K}$.

In each publication in this thesis, we assume bidders to have quasi-linear utility functions of the form: $u_i = v_i(S) - p_i$ with S being some set of items allocated to i and p_i being a monetary sum that is determined based on some payment rule (see 2.2.3) and needs to be paid by bidder i.

Typically, economists aim to maximize welfare and therefore seek to assign the items in a way that the bidders' sum of valuations are maximized. Assuming bids are a reliable indicator of bidders' valuations, in a single item auction we assign the item to the highest bidder. In an auction with multiple items and the possibility to submit a bid for any combination of items, i.e. in a combinatorial auction, we can formulate this allocation problem as an integer program, often referred to as the winner determination problem (WDP):

$$
\begin{array}{llll}
\max\limits_{x_i(S)} & \sum_{S \subseteq \mathcal{K}} \sum_{i \in \mathcal{I}} x_i(S) v_i(S) & & \\
s.t. & \sum_{S: k \in S} \sum_{i \in \mathcal{I}} x_i(S) & \leq 1 & \forall k \in \mathcal{K} \\
 & \sum_{S \subseteq \mathcal{K}} x_i(S) & \leq 1 & \forall i \in \mathcal{I} \\
 & x_i(S) & \in \{0, 1\} & \forall S \subseteq \mathcal{K}, \forall i \in \mathcal{I}
\end{array} \tag{WDP}
$$

This WDP determines the allocation of bundles to bidders, $x_i(S)$, such that it maximizes the valuations of all bidders. Here, we assume that we have access to the true valuations of bidders $v_i(S)$. In practice, we usually have no access to the true valuations and use the submitted bids as a proxy. At the same time, we ensure that each item is allocated no more than once as well as that each bidder is allocated at most one bundle.

After determining the allocation of items, a payment rule determines the monetary amount that is transferred from the winning bidders to the auctioneer.

2.2.3 Payment Rules and Properties

Before introducing the most relevant payment rules applied throughout this thesis, let us first introduce three mechanism design criteria in the context of auctions that we consider important:

The first criterion is allocation *efficiency*. An allocation is considered efficient if it maximizes welfare, i.e. the sum of the valuation of all bidders. The WDP is efficient under the assumption that the bids correctly represent a bidder's valuation.

The second criterion is *incentive-compatibility*. An auction is *incentive-compatible* if it is best for each bidder to act according to their true valuations. This solves the problem of finding an optimal strategy $\beta(v)$ since reporting truthfully is a weakly dominant strategy for each bidder.

The third criterion is that no bidder should be worse off having participated in an auction. This is referred to as *individual rationality*. If a payment rule can generally ask truthful bidders for payments larger than their received valuation, i.e. $p_i > v_i(x)$, an auction is not *individual rational* since bidders can have negative utility by participating in the auction.

In the following, we briefly introduce four payment rules, starting with two simple payment rules frequently applied in single-item auctions and continuing with two payment rules explicitly designed for combinatorial auctions. All of the presented payment rules are considered to be sealed-bid, i.e. bids are submitted without the opponents' knowledge regarding the bid amount and can subsequently not be altered. This is in contrast to some iterative auction procedures such as ascending auctions in which bidders openly submit their bids and can raise them once an opponent's bid is higher. We end with the introduction of Walrasian equilibrium prices, which are less considered as a payment rule but rather as market clearing prices, resulting from a Walrasian equilibrium. Because we apply these prices in the context of a payment rule, they are introduced in this section.

First-Price

Payments in a sealed-bid auction based on a first-price payment rule are straightforward. Here, the winning bidder pays exactly the amount that she bid on the item, i.e. $p_i = b_i$ for the single winning bidder i.

However, a bidder's optimal bid strategy in a first-price payment rule is not straightforward. If she bids too high, i.e. very close or equal to her valuation, she might win but loses margin; if she bids too low she might increase her margin but eventually she might lose because her bid is lower than the opponents' highest bid. Equilibrium strategies for simple single-item auctions exist and are described for example in Krishna (2009), p.14.

Second-Price

The second-price payment rule solves the strategic problem presented above. Here, the winner's payment is set to the second highest bid, therefore: $p_i = \max\limits_{j \neq i} b_j$, for the single winning bidder i.

The winner's payment is no longer determined by her own bid but by the opponents' bids. In a second-price auction it is a weakly dominant strategy to bid truthfully, i.e. $b_i = v_i$, $\forall i \in \mathcal{I}$. The auction is strategy-proof, i.e. dominant-strategy incentive-compatible.

Vickrey-Clarke-Groves

The second-price payment rule above can easily be applied to single-item auctions but not to combinatorial auctions. When multiple items are auctioned at once, the Vickrey-Clarke-Groves (VCG) mechanism (Vickrey, 1961) provides a general solution to achieve dominant-strategy incentive compatibility in settings with quasi-linear utility functions. The idea is that each winning bidder $i \in \mathcal{W}$, with $\mathcal{W} \subseteq \mathcal{I}$ being the set of winners, pays an amount equal to the opponents' loss caused by having i participate in the auction; or equally, each winning bidder $i \in \mathcal{W}$ pays the valuation they receive in the welfare maximizing allocation, $v_i(x^*)$, minus a discount in the additional welfare w generated by i participating in the auction. The additional welfare is given by the difference in welfare generated by the set of all bidders $\mathcal{I}$, $w(\mathcal{I})$, and the welfare generated by all bidders without i, $w(\mathcal{I}_{-i})$, therefore:

$$p_i = v_i(x^*) - (w(\mathcal{I}) - w(\mathcal{I}_{-i}))$$

The mechanism incentivises each bidder to submit their bids truthfully. It also requires each bidder to submit bids on all possible bundles, therefore an exponential amount of bids in the number of items. As a result, an enormous amount of information might be required, which is one of the reasons why VCG is barely applied in practice.

Another reason is that the resulting prices can be outside the core. For the time being, we consider the auction as a *cooperative* game $(\mathcal{N}, w)$ where $\mathcal{N}$ is the set of players, here including the auctioneer, who can form *coalitions* with transferable utilities and can freely move won items among them. w defines the coalitional value function for any coalition $\mathcal{C} \in \mathcal{N}$ as the maximum sum of the participants' valuation for an allocation alternative x out of the set of alternatives $\mathcal{X}$. The coalitional value function is:

$$w(\mathcal{C}) = \left\{ \begin{array}{ll} \max_{x \in \mathcal{X}} \left\{ \sum_{i \in \mathcal{C}} v_i\left(x_i\right) \right\}, & \text{if } 0 \in \mathcal{C} \\ 0, & \text{if } 0 \notin \mathcal{C} \end{array} \right\}$$

For the value function to be positive the auctioneer must be part of the coalition, because else the bidders have nothing to trade among themselves.

Let $u_i = v_i(x_i) - p_i$ be the profit of player i, the set of core payoffs is defined as:

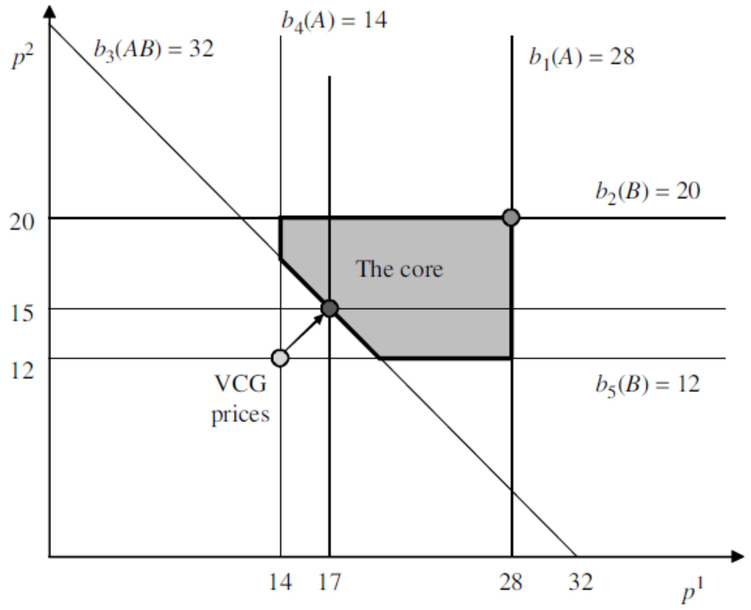


Figure 2.1: Core violating VCG payments from Example 1 in Day and Cramton (2012)

$$Core(\mathcal{N}, w) = \left\{ u | \sum_{i \in \mathcal{N}} u_i = w(\mathcal{N}), (\forall \mathcal{C} \subset \mathcal{N}) w(\mathcal{C}) \leq \sum_{i \in \mathcal{C}} u_i) \right\}$$

If some payoff vector u is not in the core, there exists a coalition $\mathcal{C}$ that in total is better off by deviating from the current allocation at current prices and sharing the total surplus. The example below illustrates this problem.

Example 2.2: Core Violating VCG Payments, (based on Day and Cramton (2012), Example 1). *Consider 5 bidders $\mathcal{I} = \{1, \cdots, 5\}$ and 2 items $\mathcal{K} = \{A, B\}$. Each bidder submits only one bid equal to her valuation. The submitted bids are the following: $b_1(A) = 28$, $b_2(B) = 20$, $b_3(AB) = 32$, $b_4(A) = 14$, $b_5(B) = 12$. The unique winners in this example are: $\mathcal{W} = \{1, 2\}$; the resulting positive VCG payments are: $p_1^{VCG} = 14$, $p_2^{VCG} = 12$. Note that both bidders together pay 26 while bidder 3 would be willing to pay 32 and currently has a payoff of 0. She could therefore form a coalition with the auctioneer and offer a payment larger than 26 but less than her valuation to receive a positive payoff.*

Core-Price

When VCG payments are outside the core (e.g. see Example 2.2), they can be increased in a way that they are in the core, i.e. any deviating coalition does no longer exist. Note that this is always possible in auctions because a pay-as-bid, i.e. first-price, payment rule is always in the core. There exist multiple different algorithms to adjust the VCG prices. One of them is the VCG-nearest core payment by Day and Raghavan (2007) and its refined version by Day and Cramton (2012). Day and Cramton (2012) compute core-payments such that the Euclidean distance to the VCG payments is minimal. Considering the example of core violating VCG payments above, the payments would be increased to $p_1^{Core} = 17$ and $p_2^{Core} = 15$ as can be seen in Figure 2.1. Now, no one would be willing to deviate from the current allocation at the given prices and the outcome is in the core. However, since core prices deviate from VCG prices it can no longer be guaranteed that the mechanism is strategy-proof. By keeping this deviation minimal, the incentive to deviate from truthful bidding is meant to be kept minimal, too.

Walrasian Equilibrium Price

Arrow and Debreu (1954) introduced Walrasian equilibrium as an equilibrium concept in markets with divisible items or where the quantity exchanged between traders is small compared to the total quantity being traded. Due to the size of the market, the individual market participant cannot manipulate the price.

In the context of combinatorial auctions, Blumrosen and Nisam (Nisan (2007), p. 275 - 279) connect linear programming with Walrasian equilibria. They prove that "if an integral optimal solution exists for the linear programming relaxation", i.e. the optimal solution of a relaxed WDP (Section 2.2.2) returns decision variables that are integer, "then a Walrasian equilibrium whose allocation is the given solution also exists. [...] A Walrasian equilibrium is a set of *market-clearing* prices where every bidder receives a bundle in her demand set, and unallocated items have zero prices" We can access these item level prices through the dual of the WDP.

In Chapter 3 (Sutterer et al., 2019) we rely on the implementation of a WDP using VCG and Core payments for the implementation of an optimal offline allocation and (nearly) incentive-compatible payment rule. For an online allocation we apply the simple second-price payment rule. In Chapter 4 (Bichler et al., 2020) we implement two additional payment rules to the ones presented here, namely discriminatory payments and uniform payments (Krishna (2009), p. 174 - 178). Because they are described in detail in Chapter 4 (Bichler et al., 2020) and are used only therein, we refrain from introducing them twice. For the implementation of our proposed combinatorial auction design, we build on the theory of Walrasian equilibrium prices presented here. In Chapter 5 (Heidekrüger et al., 2019) and Chapter 6 (Heidekrüger et al., 2020a) we compare the performance of strategies learned via our developed learning algorithm Neural Pseudogradient Ascent, to those playing the Bayes-Nash equilibrium. We consider settings with first-price, second-price, VCG and Core payment rules.

2.3 Learning in Incomplete Information Games with Continous State and Action Space

In Section 2.1.2 we have introduced Fictitious Play and Smooth Fictitious Play as learning algorithms, originating from the economics branch of research to learning games of complete information with discrete action and state space. In fact, both can also be applied to games of incomplete information as long as the state and action space remains discrete.

For the Bayesian auction games we consider in this thesis, actions as well as states, i.e. types, are continuous. In Chapter 5 (Heidekrüger et al., 2019) and 6 (Heidekrüger et al., 2020a), we rely on the implementation of fully connected neural networks to compute policy gradients in continuous state and action space, in order to learn Bayes-Nash equilibria. In Chapter 5 we refer to this learning algorithm as Neural Self-Play while in Chapter 6 we refer to it as Neural Pseudogradient Ascent (NPGA).

In the following, we explain the concept of fully connected neural networks and evolutionary strategies to provide additional background in order to elevate the understanding for the implementation of NPGA. We restrict ourselves to a brief and concise description that should provide the reader with the necessary understanding. The following introduction is mostly based on Chapter 6 in Goodfellow et al. (2016). For further information on reinforcement learning and neural networks we refer to Goodfellow et al. (2016) and Sutton and Barto (2018).

2.3.1 Fully Connected Feedforward Neural Networks

With increasing computational power and availability of mass data at the beginning of the 21st century, neural networks have become a popular method for all kinds of machine learning applications. These range from simple prediction tasks to picture classification to applications in multi-agent reinforcement learning.

Definition 2.12: Neural Network. *A neural network defines a mapping $y = f(x, \theta)$ with x being the input data and θ being parameters to be learned for the best function approximation. The network is composed of a number of sequential functions, i.e. a chain: $f(x) = f^{(n)}(f^{(n-1)}(\cdots(f^{(1)}(x))))$ that can be represented by an acyclic graph. The superscript describes the hidden layers 1 to $n-1$, ending with the output layer n. Each of these layers consists of a predefined number of neurons, representing a vector to scalar operation, and each neuron applies an activation function to allow for non-linear transformations.*

Information in the form of the input x flows through the network, with parameters θ, and finally result in the output y. This is referred to as a *forward pass* through all n layers of the network. The number of layers determines the depth and the number of *neurons* determines the width of the network.

Definition 2.13: Neuron. *A neuron represents a vector to scalar operation: $z = g(Wx + b)$ connecting the layers. x is the output data of the previous layer and z is the output of the neuron in the current layer. W and b are parameters that are part of the training process and g is referred to as an activation function.*

In a fully connected neural network each neuron of one layer is connected to each neuron of the following layer. Weights $W_r^{(l)}$ represent a weight vector of neuron r in layer l connecting neurons of the previous layer to the current. $b_r^{(l)}$ is a scalar bias value of r in l. Figure 2.2 illustrates a typical fully connected neural network with 1 hidden layer.

Each neuron applies an *activation function* to its affine transformation, $Wx + b$, before returning the output z.

Definition 2.14: Activation Function. *An activation function g transforms the affine transformation of a neuron, $Wx+b$, and allows for non-linear transformation. It is part of the parameters defining the architecture of the neural network and is not trained.*

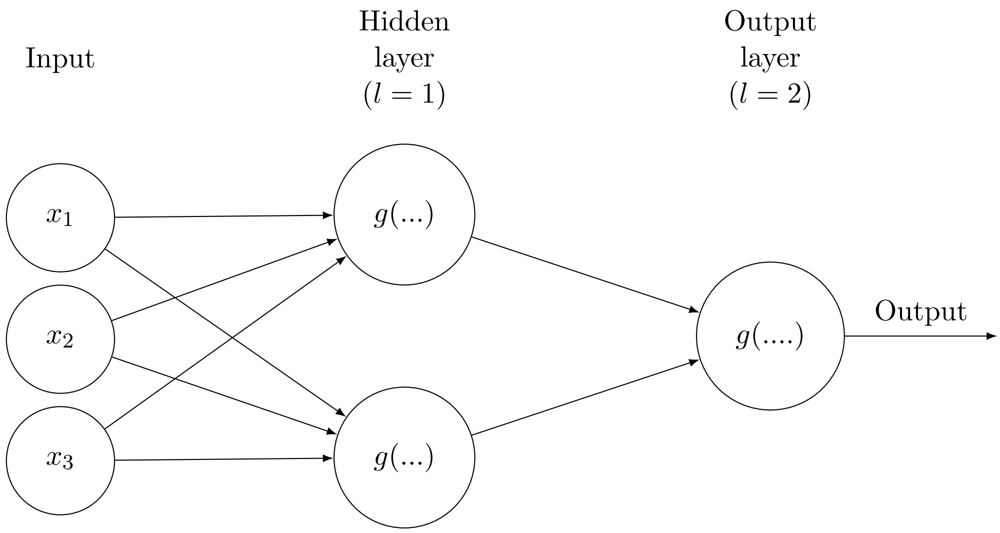


Figure 2.2: Fully Connected Neural Network

Typically, all neurons of a single layer use the same activation function. Common activation functions are: sigmoid, tanh as well as SeLU and ReLU. We use the latter two for our network design, SeLU in all hidden layers and ReLU in the output layer. They are defined as:

$$SeLU(x) = \lambda \left\{ \begin{array}{ll} x, & \text{for } x \geq 0 \\ \alpha e^x - \alpha, & \text{for } x < 0 \end{array} \right\} \quad ReLU(x) = \left\{ \begin{array}{ll} x, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{array} \right\}$$

SeLU is an approximation to ReLU that provides good numerical properties for certain values of λ and α effectively working as an internal normalization (Klambauer et al., 2017). It has become very popular lately and has in fact shown to perform well in our settings, too. Using ReLU in the output layer provides a natural way to avoid negative bids, which are not allowed in the bidding process.

Figure 2.3 illustrates the computational pipeline of a single neuron. The neuron here has 3 input connections (x_1, x_2, x_3). Each is multiplied with its own weight $(w_{r,1}^{(l)}, w_{r,2}^{(l)}, w_{r,3}^{(l)})$. The bias, $b_r^{(l)}$, is added to the summed products of input and weight. Its resulting scalar is the input of the activation function g leading to output z.

The following example illustrates the usage of a neural network in the context of auctions:

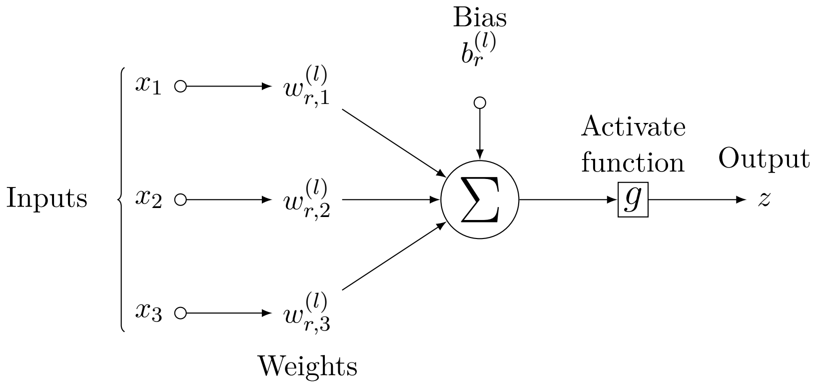


Figure 2.3: Neural Network Pipeline

Example 2.3: Bidding Strategies via Neural Networks. *Consider a first-price auction for a single item. Bidder i uses a fully connected neural network to compute her bid. The input is a scalar valuation $x = v_i$. The network has 1 hidden layer $l = 1$ followed by an output layer $l = 2$. The hidden layer has 2 neurons and the output layer has 1 neuron. All layers use a ReLU activation function. Assume the initial weights are:* $w_{1,1}^{(1)} = -1.1, w_{2,1}^{(1)} = 1.2, w_{1,1}^{(2)} = 2.1, w_{1,2}^{(2)} = 2.2$ *and* $b_1^{(1)} = b_2^{(1)} = b_1^{(2)} = -0.5$.

Bidder i has a valuation of $v_i = 1$ for the item. With $x = v_i$ the bid is computed as follows:

- *Input for neuron 1 and 2 in layer 1:* $x = 1$
- *Output of neuron 1 in layer 1:* $ReLU(-1.1 \cdot 1 - 0.5) = 0$
- *Output of neuron 2 in layer 1:* $ReLU(1.2 \cdot 1 - 0.5) = 0.7$
  - *Inputs for neuron 1 in layer 2:* $x = (z_{1,1} = 0, z_{1,2} = 0.7)$
  - *Output of neuron 1 in layer 2:* $ReLU\left((2.1 \cdot 0 + 2.2 \cdot 0.7) - 0.5\right) = 1.04$

The resulting bid is 1.04. *Assume the second highest bid is 0.7, bidder i wins the item but pays more than the item is worth to her.*

In this case bidder i would overbid. During training, the parameters θ will be adjusted such that i should bid less in the next training iteration.

2.3.2 Gradient Computation

To be able to train a neural network we need to evaluate the performance of its output first. The output in NPGA is an action, i.e. a bid, by a player, i.e. a bidder, that results in the bidder's utility u. As a consequence, we train a Loss function $L(\theta) = -\mathbb{E}[u]$. An efficient way of optimizing loss functions in neural networks with batches of data x is backpropagation. A batch of data denotes a subset of the complete data set for which one gradient can be calculated at once. Backpropagation computes the gradient of the loss function $\nabla_\theta L$, i.e. the direction of the steepest increase of the function. It can efficiently use the chain rule to compute and store the partial derivatives of each neuron.

As explained in Chapter 5 and 6, we cannot use backpropagation because of nontrivial discontinuities in a bidder's ex-post utilities. Eventually, backpropagation would lead to zero bids and as a result would lead to zero gradients that do not constitute an equilibrium. Because of the shape of the ReLU function when $x < 0$, the gradients do not receive any more signals and are stuck. This is sometimes referred to as "dead ReLU". To avoid this from happening, we apply evolutionary strategies (Salimans et al., 2017) instead.

Evolutionary Strategies

Evolutionary strategies (Salimans et al., 2017) create P random perturbations of the parameter vector θ by adding P i.i.d. zero-mean, σ^2 variance Gaussian noise terms $\varepsilon_1, ..., \varepsilon_P$. The resulting P perturbed models are evaluated with respect to their "fitness", i.e. achieved utility, $\varphi_p \in \mathbb{R}$. This is used to build the weighted average of the P noise vectors with more desirable perturbations being weighted higher than less desirable ones:

$$\nabla_\theta L = -\frac{1}{\sigma^2 P} \sum_{p=1}^{P} \varphi_p \varepsilon_p$$

Example 2.4: Pseudogradient Computation. *Consider the setting and network as described in Example 2.3. We compute $P = 2$ perturbations of the parameter vector $\theta = (w^{(1)}_{1,1}, b^{(1)}_1, w^{(1)}_{2,1}, b^{(1)}_2, w^{(2)}_{1,1}, w^{(2)}_{1,2}, b^{(2)}_1)$. Let us assume $\varepsilon_1 = (0.4, 0.2, -0.3, 0.1, -0.3, 0.2, 0.1)$ and $\varepsilon_2 = (-0.2, -0.2, 0.3, 0.3, 0.2, 0.1, -0.2)$. The corresponding parameters for the* *first perturbation change to:* $w_{1,1}^{(1)} = -0.7, b_1^{(1)} = -0.3, w_{2,1}^{(1)} = 0.9, b_2^{(1)} = -0.4, w_{1,1}^{(2)} = 1.8, w_{1,2}^{(2)} = 2.4, b_1^{(2)} = -0.4$. *The resulting bid for* $\theta + \varepsilon_1$ *is:*

$$ReLU\left((1.8 \cdot 0 + 2.4 \cdot 0.5) - 0.4\right) = 0.8$$

making the bidder win at a price of 0.8 with a payoff $u_i = 0.2$. *The second perturbation,* $\theta + \varepsilon_2$, *leads to a bid of* 2.29 *with a payoff* $u_i = -1.29$. *As a consequence* $\varphi_1 > \varphi_2$ *and the actual parameter vector* θ *is updated more in the direction of* ε_1 *and less in that of* ε_2, *effectively decreasing the bidder's bid in the next iteration.*

2.3.3 Optimization

Backpropagation and evolutionary strategies, provide a (pseudo) gradient that merely determines the direction for optimization but not the step size. Finding a good step size represents a whole new optimization problem. The machine learning community has presented many good working approaches, starting with Stochastic Gradient Descent to Momentum (Qian, 1999) and Adam (Kingma and Ba, 2017). For our work, we tried all three optimization algorithms. Depending on the setting, they sometimes perform better or worse. We finally chose Adam as it produces the most robust results. With Adam, the parameter update for the next iteration $k+1$ is computed as:

$$\theta^{k+1} = \theta^k - \alpha \cdot \frac{m^{k+1}}{\sqrt{v^{k+1}} + \epsilon}$$

using the first momentum, the mean of the gradients:

$$m^{k+1} = \beta_1 \cdot m^k + (1 - \beta_1)\nabla_\theta L(\theta^k)$$

and the second momentum, the variance of the gradients:

$$v^{k+1} = \beta_2 \cdot v^k + (1 - \beta_2)[\nabla_\theta L(\theta^k) \circ \nabla_\theta L(\theta^k)],$$

$\circ$ being an element-wise multiplication. α, β_1, β_2 and ϵ are all hyperparameters that need to be determined a priori and can be tuned.

After updating the parameter vector θ, the next iteration starts with another forward pass, gradient computation via evolutionary strategies, parameter update via Adam, etc. Typically, we perform multiple thousand iterations with large batch sizes in our experiments.

After having briefly introduced fundamentals in game and auction theory to elevate the understanding for the methodologies applied in this thesis' publications, we continue with presenting those publications. In Chapter 7 we review our publications and emphasize the contribution of each to the research community and practitioners. Finally, we outline the conclusion of this dissertation in Chapter 8.

3 Efficiency and Revenue in Display Ad Auctions

Peer-Reviewed Journal Paper

Title: Are Truthful Bidders Paying too Much: Efficiency and Revenue in Display Ad Auctions

Authors: P. Sutterer, S. Waldherr, M. Bichler

In: ACM Transactions on Management Information Systems

Abstract: Display ad auctions have become the predominant means to allocate user impressions on a website to advertisers. These auctions are conducted in milliseconds online, whenever a user visits a website. The impressions are typically priced via a simple second-price rule. For single-item auctions, this Vickrey payment rule is known to be incentive-compatible. However, it is unclear whether bidders should still bid truthful in an online auction where impressions (or items) arrive dynamically over time and their valuations are not separable as is the case with campaign targets or budgets. The allocation process might not maximize welfare and the payments can differ substantially from those paid in an offline auction with a Vickrey-Clarke-Groves payment rule, or also competitive equilibrium prices. We study the properties of the offline problem and model it as a mathematical program. In numerical experiments, we find that the welfare achieved in the online auction process with truthful bidders is high compared to the theoretical worst case efficiency, but that the bidders pay significantly more on average compared to what they would need to pay in a corresponding offline auction in thin markets with up to four bidders. However, incentives for bid shading in these second-price auctions decrease quickly with additional competition and bidders risk losing.

Are Truthful Bidders Paying too Much? Efficiency and Revenue in Display Ad Auctions

PAUL SUTTERER, STEFAN WALDHERR, and MARTIN BICHLER,
Technical University of Munich, Germany

6

Display ad auctions have become the predominant means to allocate user impressions on a website to advertisers. These auctions are conducted in milliseconds online, whenever a user visits a website. The impressions are typically priced via a simple second-price rule. For single-item auctions, this Vickrey payment rule is known to be incentive-compatible. However, it is unclear whether bidders should still bid truthful in an online auction where impressions (or items) arrive dynamically over time and their valuations are not separable, as is the case with campaign targets or budgets. The allocation process might not maximize welfare and the payments can differ substantially from those paid in an offline auction with a Vickrey-Clarke-Groves (VCG) payment rule or also competitive equilibrium prices. We study the properties of the offline problem and model it as a mathematical program. In numerical experiments, we find that the welfare achieved in the online auction process with truthful bidders is high compared to the theoretical worst-case efficiency, but that the bidders pay significantly more on average compared to what they would need to pay in a corresponding offline auction in thin markets with up to four bidders. However, incentives for bid shading in these second-price auctions decrease quickly with additional competition and bidders risk losing.

CCS Concepts: • **Information systems** → **Display advertising**; *Online auctions*; • **Computing methodologies** → *Optimization algorithms*; *Discrete-event simulation*; • **Social and professional topics** → Online auctions policy; • **Applied computing** → Marketing;

Additional Key Words and Phrases: Display ad auctions, efficiency, incentives in online auctions, real-time bidding

ACM Reference format:
Paul Sutterer, Stefan Waldherr, and Martin Bichler. 2019. Are Truthful Bidders Paying too Much? Efficiency and Revenue in Display Ad Auctions. *ACM Trans. Manage. Inf. Syst.* 10, 2, Article 6 (June 2019), 18 pages.
https://doi.org/10.1145/3325523

1 INTRODUCTION

Online advertising markets like keyword auctions on major search engines have developed into an increasingly popular topic in information systems research (Chen and Stallaert 2014; De et al. 2010; Yang and Ghose 2010; Zhao and Xue 2012). Apart from keyword auctions, the display ad auctions have become very important. The market has grown to represent more than one-third

Authors' addresses: P. Sutterer (corresponding author), Technical University of Munich, Boltzmannstrasse 3, Garching/Munich, Bavaria, 85748, Germany; email: paul.sutterer@tum.de; S. Waldherr and M. Bichler, Technical University of Munich, Garching, Germany; emails: {stefan.waldherr, bichler}@in.tum.de.

2158-656X/2019/06-ART6 $15.00
https://doi.org/10.1145/3325523

of total media spending, expected to be surpassing spending on TV advertisement in the U.S. in 2017 (Zenith 2017). Display ad auctions are conducted in milliseconds online, whenever a user visits a website. This is also referred to as real-time bidding (RTB). Based on available information about the user and the publisher hosting the website, advertisers determine their valuation for an impression and place a bid.

The impressions are typically priced via a simple second-price rule (Yuan et al. 2013). For single-item auctions, this Vickrey payment rule is known to be incentive-compatible (Vickrey 1961), i.e., bidders maximize their profit by reporting truthfully. Economic models of keyword and display ad auctions often assume that the valuations for individual impressions are separable, i.e., a bidder's valuation of each impression does not depend on the allocation of other impressions. Each auction can be considered independently such that the second-price rule is indeed dominant-strategy incentive-compatible (see, for example, the assumptions in a widely cited paper by Edelman and Schwarz (2010)). The situation is quite different if the valuations for individual impressions are not separable anymore. This might be the case, if there is an overall budget constraint or a campaign target, i.e., a budget beyond which bidders do not have a value for additional impressions anymore. If all impressions would be auctioned offline at the same time with a Vickrey-Clarke-Groves (VCG) mechanism, then this mechanism would still be incentive compatible for bidders with a campaign target. This, however, is not necessarily the case if the impressions are auctioned off dynamically. With non-separable preferences a sequential auction does not necessarily pick the welfare maximizing allocation. However, optimality of the allocation problem is a requirement for the VCG mechanism to be truthful (Mas-Colell et al. 1995).

An online display ad auction (with truthful bidders) can be seen as an *online algorithm* to solve a computationally hard knapsack problem. Unfortunately, worst-case analysis yields that the efficiency of online algorithms can be arbitrarily low for many knapsack problems (Berg et al. 2010; Chakrabarty et al. 2008). In summary, the online auction might not maximize welfare, and even if bidders bid truthfully, the payments can differ substantially from those paid in an offline auction with a VCG payment rule or also competitive equilibrium prices. Equilibrium bidding strategies in such dynamic auctions are not well understood and require strong assumptions about the nature of the arrival process, the campaign length, and the utility functions (Balseiro et al. 2015). It is unclear whether advertisers follow such strategies in the field. Given the second-price rule, a simple strategy, where advertisers express their value for an impression truthfully, is easy to implement and it might still turn out to be a good strategy with high payoff for the bidders. We want to understand under which circumstances we can expect this to be the case.

1.1 Contributions and Outline

In this article, we analyze the allocative efficiency and the bidders' payoff in online display ad auctions when bidders report truthfully. For this, we first study the allocation and prices of the offline allocation problem, which serves as a baseline to compare the outcome of an online display ad auction against. This means, we optimally allocate impressions and compute prices in a hypothetical market, where all impressions arrive at the same time.

The offline allocation problem is computationally hard and results in a binary integer program. We compute Vickrey-Clarke-Grove (VCG) payments, which are dominant-strategy incentive-compatible, also for non-separable valuations. Unfortunately, in multi-object auctions such prices might differ from competitive equilibrium (aka. core) prices unless the VCG outcome is in the core. In a first theoretical contribution, we show that with a typical value model for display ad auctions, the outcome of VCG auction is not guaranteed to be in the core. We compute bidder-optimal core payments and argue that, while VCG payments might be very low, core payments provide a reasonable comparison for prices achieved in an online display ad auction. In a second theoretical contribution, we show that the efficiency of an online display ad auction might be as low as 50% in the worst case.

Finally, we study a simple experimental market to better understand the average case efficiency and prices in online display ad auction when bidders would report truthfully. In these numerical experiments, we find that the welfare achieved in the online display ad auction with truthful bidders is very high, and the worst-case analysis is too pessimistic as an estimator for the efficiency of such auctions in the field.

However, we also show that the bidders pay substantially more on average compared to what they would need to pay in a corresponding offline auction in which all impressions are available at one point in time. This is due to the fact that truthful bidders bid on more impressions in the online process than in the offline problem, which drives up prices. This sets incentives for demand reduction and non truthful reporting of only a fraction of the actual bidder's valuation (bid shading). Indeed, such bidding strategies have also been reported in the field (Balseiro et al. 2015; Zhang et al. 2014). Therefore, we also consider the case when one bidder unilaterally deviates from truthful bidding by shading his bid. Our results indicate that bidders have incentives to deviate from truthful bidding only in thin markets with few competitors. With more competitors, bidders lose on average when shading their bids. This is important for auctioneers and bidders in the field to understand.

The article is structured as follows. First, we define our economic model more formally. Afterwards, we address the allocation and pricing problem in the offline case, followed by a discussion of the online model. In the subsequent sections, we describe the experimental design and evaluate the results of the numerical experiments. Finally, we provide a summary and conclusion.

2 ECONOMIC MODEL

Advertisers in online display ad markets typically want to fulfill a marketing campaign, in which they are given a fixed budget and want to maximize their payoff by buying impressions with maximal total value. We introduce a simple value model for a payoff-maximizing advertiser. The values for a particular type of impression are additive up to a certain target, i.e., a dollar amount that is determined by the budget of the advertiser. Impressions beyond this campaign target have zero marginal valuation.

More formally, the participants in the display ad auction are one auctioneer and a number of $|I|$ bidders. The risk neutral auctioneer sells a set of $K = \{1, \ldots, |K|\}$ impressions of types $T = \{1, \ldots, |T|\}$. In real-world markets, $|K|$ is a very large number. For each bidder $i \in I$, B_i denotes his campaign target, and for each type $t \in T$, his valuation is v_{it}. The campaign target B_i defines a level at which, once the gained value reaches this level, the bidder has no further valuation for any type of impressions. His valuations are additive up to that campaign target, B_i. Thus, his valuation function $v(\cdot)$ takes the following form:

$$v_i(x) = \min\left\{\sum_k v_{it(k)} x_{ik}, B_i\right\}, \tag{1}$$

with $v_{it(k)}$ being the valuation of bidder i for impression k (of type $t(k)$), and x_{ik} being 1 if bidder i receives impression k and 0 else. Bidders are risk neutral and want to maximize their payoff. Let p_{ik} be the price that bidder i has to pay for impression k, then his utility function is

$$u_i(x) = v_i(x) - \sum_k p_{ik} x_{ik}. \tag{2}$$

Note that $u_i(x)$ is a traditional quasi-linear utility function where bidders are pure payoff maximizers. With this traditional quasi-linear utility function, the VCG mechanism is the unique dominant-strategy incentive-compatible mechanism (Mas-Colell et al. 1995).

Note that the campaign target in our article could also be modeled as a budget constraint on top of a quasi-linear utility function. However, the consequences of this alternative model are far from trivial. It is well-known that in the presence of private budget constraints there is no incentive-compatible mechanism (Dobzinski et al. 2008). Neither the VCG mechanism nor a core-selecting auction would be incentive-compatible and there is little we know about equilibrium strategies in markets with private budget constraints at this point. Our value model $v_i(x)$ can be seen as a reasonable approximation of how advertisers value impressions for an advertising campaign and it allows us to use VCG and core payments as benchmark to compare against.

Let us now discuss allocation and pricing in a hypothetical offline market and in an online market modeled after display ad auctions in the field.

3 ALLOCATION AND PRICING IN AN OFFLINE MARKET

The economic model now allows us to formulate the allocation and pricing problem. As usual, the allocation problem (AP) maximizes overall social welfare. For each advertiser $i \in I$ and each impression k of type $t(k) \in T$, the integer variable x_{ik} indicates whether impression k is assigned to advertiser i and b_i denotes the aggregated valuation of the impressions that advertiser i obtains.

$$\max \sum_{i \in I} b_i \tag{\textbf{AP}}$$

s.t.

$$b_i \leq \sum_{k \in K} v_{it(k)} x_{ik} \qquad \forall i \in I \tag{VL}$$

$$b_i \leq B_i \qquad \forall i \in I \tag{CL}$$

$$\sum_{i \in I} x_{ik} \leq 1 \qquad \forall k \in K \tag{Supply}$$

$$x_{ik} \in \{0, 1\} \qquad \forall k \in K, i \in I \tag{Binary}$$

For each advertiser $i \in I$, constraint (VL) states that his aggregated valuation b_i can not be higher than the sum of the valuations for each individual impression he obtains. Constraint (CL) represents the campaign target and limits the value of very large packages to that of the advertisers campaign target. Together, these constraints model the utility function of each advertiser. Constraint (Supply) assures that each impression $k \in K$ is allocated only once.

Based on this allocation, the payments for all advertisers need to be computed. In the Appendix, we show that the VCG outcome does not necessarily lie in the core.

Proposition 1. *VCG payments do not always lie in the core when considering valuation function $v(\cdot)$.*

Proof. See the Appendix. □

Since the VCG payments do not necessarily lie in the core, we compute bidder-optimal core prices to compare the payments in the online auction against. Those prices are based on VCG payments that are computed for each bidder $i \in I$ by $p_i^{VCG} = v_i(x_i^*) - [w(I) - w(I_{-i})]$, with x_i^* being the bundle allocated to i in an efficient allocation x^* and $w(I)$ being the total welfare including all bidders while $w(I_{-i})$ is the total welfare excluding bidder i. When VCG payments are not in the core, the algorithm by Day and Raghavan (2007) minimally increases the prices until there exists no coalition which would want to deviate from the allocation at current prices (bidder-optimal core prices). Such prices constitute a competitive equilibrium in the offline market, where no set of bidders would want to deviate. Note that the core is not empty, since charging all winners their true valuation results in a core imputation. Due to space restrictions, we refer the reader to Day and Raghavan (2007) for details of the core computation. While core payments do not need to coincide with VCG payments, in our simulations both payments overlap most of the time such that we can assume truthful bidding in the offline model.

4 ALLOCATION AND PRICING IN ONLINE DISPLAY AD AUCTIONS

In online display ad auctions impressions are sold sequentially within milliseconds. A second-price auction per impression is typically used to price these impressions. The highest bidder obtains the auctioned impression and pays the second-highest bid as a price to the auctioneer. Formally this is equivalent to the VCG price computation, if the valuations of bidders were separable across items. As we discussed earlier, this assumptions is typically violated in advertising markets where advertisers run campaigns for which they have a fixed budget.

Let us provide a brief example to illustrate the differences between the offline and the online allocation. Consider two bidders $I = \{b1, b2\}$ and two impressions where the first impression is of type $t(1) = 1$ and the second is of type $t(2) = 2$. Both bidders have a valuation for each impression type of $v_{it(k)} = 2, \forall i \in \{b1, b2\}, t(k) \in \{1, 2\}$; both bidders have a campaign target of $B_i = 3$. When solving the offline auction, each bidder receives one impression (which one is irrelevant, because the valuations are identical). Both bidders value their bundle of a single impression by $b_i = 2$ and the total welfare is $\sum_{i \in I} b_i = 4$. The VCG payment for a bidder i is computed by the bidders valuation for his bundle b_i minus the Vickrey discount, being the total welfare when i participates in the auction minus the total welfare when i does not participate. This results in $p_i = 2 - (4 - 3)$, since the welfare without i is 3 due to only one bidder being left and the valuation being limited by the campaign target of $B_i = 3$, for each bidder. In the online auction, the first impression is assigned randomly, since both bidders equally bid their valuation of 2. Assuming $b1$ wins, he would have to pay a price of $p_{1,1} = 2$, because this is the second highest bid. For the next impression, $b1$ would only bid 1 as his campaign target is $B_1 = 3$ and he already gained a valuation of 2. The other bidder $b2$ wins the auction by bidding his true valuation of 2 and is asked a price of $p_{2,2} = 1$ as this is the second highest bid.

In the following, we compare the allocation that can be achieved in the online setting to an optimal allocation determined by the offline model. For the efficiency of the allocation, we can prove that the allocation reached in the online scenario is at least half as good as that of an optimal allocation. This is in stark contrast to the model in which each advertiser has a fixed budget to invest in a campaign. This resembles an (online) multiple knapsack problem for which Marchetti-Spaccamela and Vercellis (1995) show that there is no non-trivial competitive ratio and the result of the online algorithm can be substantially worse than the result of an optimal offline solution.

Proposition 2. *In the worst case, the social welfare obtained in the online display ad auction with value function $v_i(x)$ is at least half the welfare achieved in an optimal offline allocation.*

Proof. See the Appendix. □

5 RESEARCH DESIGN

The worst-case analysis for efficiency, provided in the last section, might not be a good estimator for the average efficiency of online display ad auctions in the field. We use numerical experiments to better understand efficiency and pricing in such auctions and compare the results to an optimal offline solution. We are especially interested in the distribution of welfare among the seller and bidders and the possible benefits for participants by deviating from truthful bidding. In the

Table 1. Experimental Design

Parameter		All Experiments	
Number of impressions $\|K\|$		1,000	
Number of bidders $\|I\|$		3,...,15	
Bidders campaign target B_i		40,000	
Valuations v_{it}		$\Gamma(4.0, 26.0)$	
Number of impression types $\|T\|$		2,...,10	
Parameter	First Setup	Second Setup	Third Setup
Arrival rates	uniform, clustered	uniform	uniform
Shading factor	None	None	0.9,0.8,0.7,0.6,0.5,0.25

following, the properties of problem instances, the assumptions and the simulation process are described.

A single problem instance contains bidders and impressions. Impressions are defined by the following parameters: number of impression types $|T|$ and number of impressions $|K|$. All impressions $k \in K$ are sequentially created; the probability of an impression k to be of impression type $t \in T$ is given by some probability $p(t(k))$. Bidders are defined by the following parameters: number of bidders $|I|$, campaign target B_i, and valuations for each impression type v_{it}. We use a gamma distribution with parameter k_t for shape and θ for scale. These parameters are set to $\Gamma(4.0, 26.0)$, which models the distribution of prices in display ad auctions by Zhu et al. (2017). Following Zhu et al. (2017), $\theta = 26$ stays constant, while k_t changes for each impression type. The bid data generated for the experiments is available upon request.

In the offline model, the efficient allocation is obtained by solving (**AP**). Afterwards, payments are calculated using the core constraint generation algorithm (Day and Raghavan 2007), determining the bidders' and auctioneer's share of the welfare.

In the online model each impression is auctioned off sequentially with $k = 1, \ldots, |K|$. In each auction, bidders are given the impression type t and asked to submit a bid. Since we assume truthful bidding and campaign targets, a bidders bid for an impression of type $t \in T$ is given by $b_{it(k)} = \min \left\{ v_{it(k)}, B_i - \sum_{l<k} v_{it(l)} x_{il} \right\}$ where x_{il} for $l < k$ describes whether bidder i was assigned impression l. Prices are determined via the second-price rule.

We study three different experimental setups. In all setups, we consider 1,000 impressions to be auctioned off and vary the number of bidders participating between 3 and 15. This allows us to analyze several demand levels. We consider this wide range of bidders, since there is a significant variance in the number of bidders per impression in the field, however, we are not aware of publicly available statistics. For each bidder, we assume a campaign target B_i of 40,000, which makes comparisons across treatment combinations easier. The ratio between a bidders campaign target and the average valuation for one impression determines the number of impressions demanded by the bidder. The campaign target is set such that considering the current ratio and three bidders, the average demand of all bidders is more than 1,000 impressions. With this VCG prices in the offline market are strictly greater than zero on average. A summary for the parameter settings is displayed in Table 1.

Since the arrival rate of different types of impressions might not be constant and uniformly distributed across time, we analyze the arrival rate's impact on efficiency in a first experimental setup. For this, we compare uniform arrival rates with a setting where all impressions of one type arrive in one cluster (i.e., in direct succession), followed by impressions of the next type in the next cluster. For example, in the setting with a constant uniformly distributed arrival rate, the probabilities of impressions $k \in K$ to be of type $t \in \{1, 2\}$ are drawn from a uniform distribution with $p(t(k) = 1) = p(t(k) = 2) = \frac{1}{2}$ at all times. In the second setting, the probability of an impression to be of the first type is 1 in the first 500 impressions and zero later, and vice versa for impressions of type 2.

For this experimental setup, we have drawn 50 valuations randomly for each number of bidders and evaluated those on 5 randomly drawn sequences of impressions based on a uniform distribution and one sequence based on clustered arrivals. Considering 9 different numbers of impression types (2–10), we analyzed 50 (valuations) · 13 (number of bidder settings) · 9 (impression types) · (5 + 1) (impressions) = 35,100 individual experiments.

For the second and the third setup, we focus on constant uniformly distributed arrival rates. Impression types $t \in T$ for impressions $k \in K$ are drawn randomly with $p(t(k) = t) = \frac{1}{|T|}$ at all times. We vary the demand by increasing the number of bidders to study the effect on efficiency and bidder payoff (i.e., utility). We have drawn 50 valuations randomly for each number of bidders and evaluated those on 5 randomly drawn sequences of impressions based on the corresponding distribution. Thus, we analyze $50 \cdot 5 \cdot 13 \cdot 9$ (number of impression type settings) $= 29,250$ individual experiments.

While in the first and second experimental setup all bidders bid truthfully, in the third experimental setup, we consider one bidder deviating from truthful bidding. In the absence of predictive equilibrium strategies, we want to study the gains from deviating from a truthful strategy. If these gains are low, then deviations become less likely. For each problem instance, we consider each single bidder deviating from truthful bidding once, while all others continue bidding truthfully. That other bidder shades his bid with a fixed shading factor of $\lambda = \{0.9, 0.8, 0.7, 0.6, 0.5, 0.25\}$. The shading factor is multiplied with the valuation to determine the bid. For example, with $\lambda = 0.6$ and a valuation of $v_{it} = 1$ the resulting bid is 0.6. Considering 6 different levels for the shading factor, we analyze $(50 \cdot 5 \cdot 9 \cdot 13) \cdot (\sum_{i=3}^{15}(i + 1))$ (bidders deviating) · 6 (shading factors) $= 1,755,000$ individual experiments. However, these could be solved quickly, since the offline solution does not have to be recomputed.

All experiments were performed on a Xenon E3-1505M v5 @2.80GHz processor with 16GB main memory. The integer programs for optimal allocation and pricing in the offline problem are solved with Gurobi version 7.0.1.

6 RESULTS

In the following, we study the efficiency (E) and the total utility of bidders (U) by comparing online to offline welfare (W). The offline welfare (W(of)) corresponds to the optimal welfare. The efficiency loss (L) is the difference between the online welfare and the welfare in the optimal offline solution (W(of) − W(on)). The ratio between bidders utility (U) and total welfare (W) describes the distribution of welfare between auctioneer and bidders. Considering the ratio between the bidders utility (U) online to offline, we learn about the relative overpayment by bidders in the online market.

6.1 Allocative Efficiency

First, we consider varying arrival rates of impression types t. Results of the impact of different distributions are shown in Table 2. For each realized valuation set, we averaged the efficiency over five sequences of impressions and considered the average:

$$E_{avg} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{M}\sum_{j=1}^{M} E\,(\text{valuations}|\text{sequence of impressions}),$$

Table 2. Allocative Efficiency

Number of Bidders	3	4	5	6	8	10	12	15
$E_{min}(Clustered)$	0.78	0.84	0.85	0.87	0.90	0.91	0.89	0.89
$E_{min}(Uniform)$	0.88	0.88	0.89	0.88	0.92	0.91	0.93	0.94
$E_{avg}(Clustered)$	0.99	0.97	0.98	0.98	0.99	0.99	0.99	0.99
$E_{avg}(Uniform)$	0.98	0.97	0.97	0.98	0.98	0.99	0.99	0.99

where N is the number of valuation sets, M is the number of impression sets, and E(valuations| sequence of impressions) is the efficiency for a specific valuation set given a certain impression set. We also report the minimum efficiency observed for any valuation set in any of the five different sequences of impression types as E_{min}.

Result 1. *The average efficiency of the online display ad auction is close to 100% and increases with an increasing number of bidders. Even when impressions of common types arrive in clusters, the average efficiency is not significantly lower. The minimum observed efficiency is much lower but also increases with the number of bidders.*

Efficiency in our experiments is surprisingly high even though the worst-case efficiency can be as low as 50%, as we have shown formally in the last section. Although the competition for impressions in the online market is high, the impressions are always allocated to the highest valuing bidder. Inefficiencies occur when a bidder dominates all other bidders and the difference in valuation is low in the first impression type appearing while it is high for impression types appearing thereafter. The dominating bidder will win the first impressions and consume his budget before he can win the later impression types for which his value would be much higher compared to the other bidders. Chances for such extreme cases are low. For a general analysis of worst-case scenarios, we refer to Proposition 2.

When impression types arrive in clusters, the minimum observed efficiency for each setting is indeed lower, but not as low as in the worst case. An increasing number of bidders increases both, the average efficiency as well as the minimal observed efficiency. An increasing number of impression types does not affect the efficiency significantly (see Tables 9 and 10).

6.2 Welfare Distribution

Result 2. *With a large number of bidders most of the welfare gains go to the seller. In contrast, with a small number of bidders more than one quarter of the welfare gains remain with the bidders with significant differences between the online and offline auction environment.*

Table 3 describes the welfare distribution, i.e., how much of the welfare generated goes to the bidders and how much to the seller. In this table, we average over all number of impression types.

Considering a small number of bidders, the bidders' welfare share in the online market is only 40% and 31%, while in the offline market it is 94% and 66% ($\frac{U(on)}{W(on)}$ and $\frac{U(of)}{W(of)}$) in Table 3). The differences between the utility in the online and the offline auction with only a few bidders suggests that they pay too much in the online auction.

With an increasing number of bidders, the bidders' welfare share becomes very small in online and offline markets. Considering a large number of bidders, it is only about 16% in the online and 20% in the offline market (see Table 3). The offline utility is much more sensitive to competition, since VCG payments consider the valuations of all bidders for all impressions while the simple second price auction considers only the valuation of the second highest bidder for the current

Table 3. Online and Offline Results w.r.t. the Number of Bidders Participating

Number of Bidders	3	4	5	6	8	10	12	15
W	116,596	140,391	151,067	162,183	173,717	180,109	189,364	197,663
U(of)	110,517	94,950	56,906	51,378	44,823	41,657	41,364	39,322
U(on)	46,149	42,485	37,427	37,201	34,189	33,068	33,152	32,642
L	1,902	4,689	4,340	3,429	2,642	2,120	1,902	1,520
E	0.98	0.97	0.97	0.98	0.99	0.99	0.99	0.99
$\frac{U(on)}{W(on)}$	0.40	0.31	0.25	0.23	0.20	0.18	0.17	0.16
$\frac{U(of)}{W(of)}$	0.94	0.66	0.37	0.31	0.26	0.23	0.22	0.20
$\frac{U(on)}{U(of)}$	0.43	0.50	0.69	0.75	0.78	0.81	0.81	0.84

Table 4. Relative Utility (Online/Offline)

Number	Number of Impression Types								
of Bidders	2	3	4	5	6	7	8	9	10
3	0.49	0.45	0.45	0.41	0.42	0.41	0.41	0.41	0.40
4	0.58	0.61	0.52	0.52	0.50	0.48	0.43	0.44	0.40
5	0.66	0.68	0.67	0.68	0.70	0.68	0.71	0.71	0.72
6	0.67	0.70	0.71	0.76	0.73	0.77	0.78	0.77	0.80
8	0.60	0.69	0.78	0.81	0.82	0.83	0.80	0.85	0.85
10	0.66	0.72	0.79	0.80	0.86	0.83	0.86	0.87	0.88
12	0.63	0.73	0.78	0.82	0.82	0.87	0.87	0.89	0.90
15	0.67	0.73	0.81	0.87	0.85	0.87	0.89	0.91	0.91

impression when determining the price. The fact that the differences between offline and online markets are small in markets with many competitors suggests incentives for deviation from truthful bidding are small in the online treatment. In Section 6.4, we analyze the gains bidders can achieve by deviating from truthful bidding. This can be seen as a measure for robustness against strategic manipulation.

6.3 Bidders' Utility in Online vs. Offline Markets

Let us now take a closer look at the differences in bidders' utility between online and offline markets taking into account also the number of impressions types that are available. These can have an impact, because bidders have different preferences for different types of impressions leading to a further segregation of the overall market and the competition for an individual impression decreases.

Result 3. *With a small number of bidders and impression types the bidders' utility from the online display ad auction is significantly lower than that of the offline auction. With an increasing number of bidders and impression types the differences between both mechanisms become small.*

Table 4 shows that with a medium number of bidders the ratio of the bidder utility in the online auction compared to the offline auction ($\frac{U(on)}{U(of)}$) increases with the number of impression types in the market. With many bidders and many impression types the utility of bidders in the online

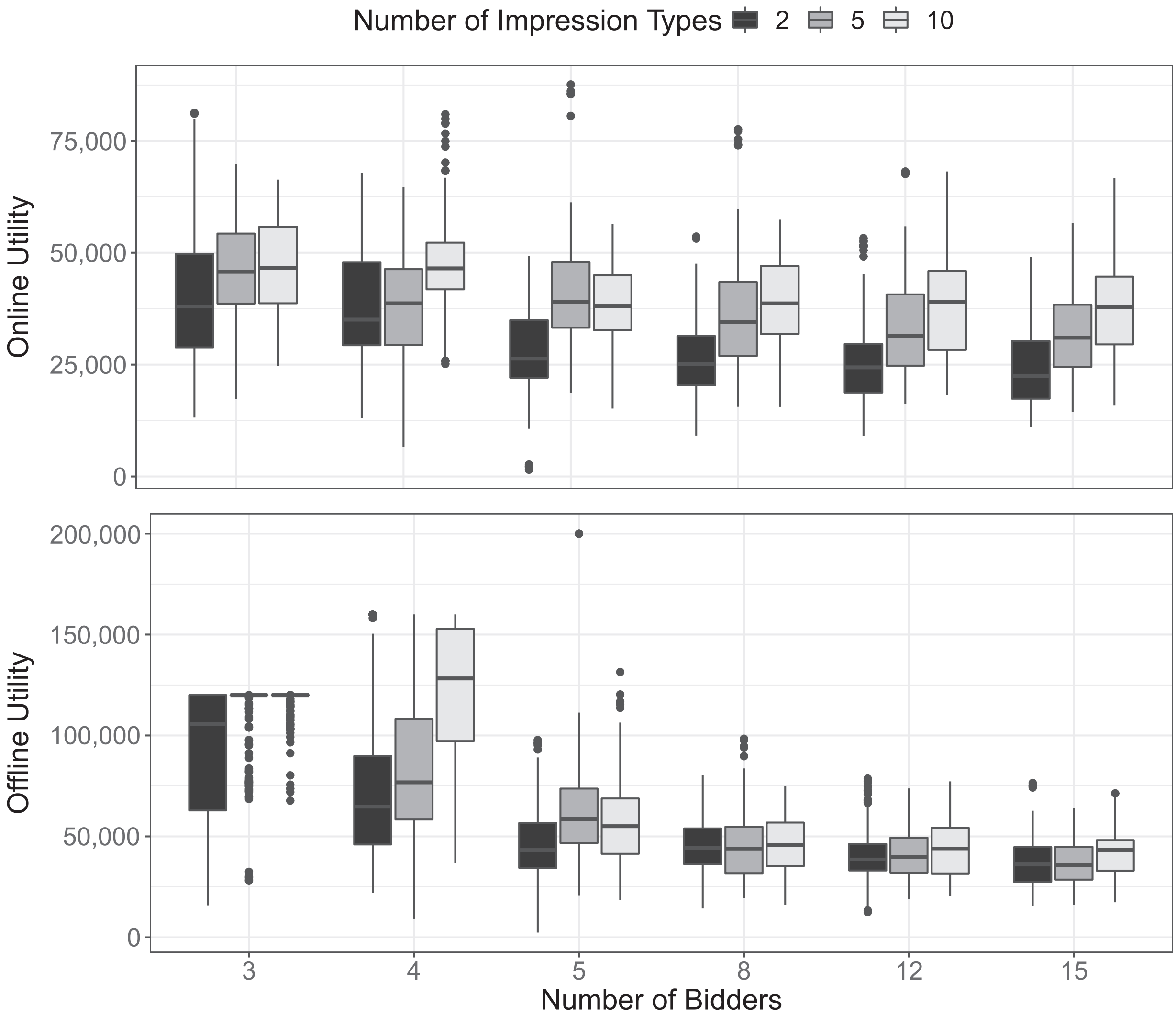


Fig. 1. Online and offline utility w.r.t. the number of impression types.

auction is close to that in the offline auction (a ratio of 0.91), while there is a substantial difference (a ratio of only 0.49) in markets with only few impression types and few bidders.

Note that with a small number of bidders an increasing number of impression types does not lead to higher utility for bidders. This phenomenon can be explained when looking at the number of impression types in online and offline markets separately (see Figure 1). A higher number of impression types further separates the market as bidders have also different preferences for different impression types even though the overall number of impressions does not change. As a consequence, the bidder utility increases with additional impression types, while it decreases with additional bidders. Note that the maximum total utility with three bidders is $3 \cdot 40{,}000(\text{budget}) = 120{,}000$, which happens often when VCG prices are 0 because of extremely low competition.

In the offline auction the bidders' utility also decreases with an increasing number of bidders. However, with a medium number of bidders there is no significant difference in the utility with two, five, or ten impression types. In contrast, with a small number of bidders, a larger number of impression types can lead to situations in the offline auction, where each of the few bidders only bids on his most preferred impression types and there is almost no competition. This leads to low prices and high utility with only a few bidders and an increasing number of impression types. With more bidders, these utility gains are quickly eaten up.

Table 7 in the Appendix summarizes the total utility of bidders in the online auction, while Table 8 shows the bidders' utility in the offline auction in numbers.

In summary, in display ad markets where the number of bidders per impression is low, the bidders pay significantly more in the online auction compared to the hypothetical offline market. As a consequence, there might be incentives for bid shading. However, considering a medium or large number of bidders and many impression types the differences decrease rapidly (also see linear regression in Table 6 in the Appendix).

6.4 Gains from Non-truthful Bidding

The differences in the utility between online and offline auctions suggests that bidders might deviate from truthful bidding in our model where bidders have a campaign target. This raises the question how bidders would behave in such a market and when they would deviate from truthful bidding. Equilibrium analysis is how economists typically aim to predict bidder behavior in markets. Auction theory has largely been focusing on offline environments where demand and supply are present at one point in time (Krishna 2009). The study of online markets where supply arrives over time is much less developed and even the development of appropriate equilibrium solution concepts is in its infancy (Balseiro et al. 2015). But even in the offline environment equilibrium bidding strategies are only known for very specific multi-object auction mechanisms. Given the number of assumptions that such equilibrium concepts require, it is also an open question whether they will be predictive for real-world markets with many objects and bidders.

Still, the second-price payment rule used in display ad auctions might set sufficient incentives to bid truthful as bidders do not have to pay their bid. Whether this is the case also depends on the expected gains a bidder has from shading his bid below his valuation in a sequence of second-price auctions. If these gains are low and the risks of losing are high, then bidders have little incentives to manipulate. In contrast, if the gains are high, this could set incentives to manipulate bids. Indeed, there are publications that recommend bid shading (Balseiro et al. 2015; Zhang et al. 2014). Note that the model assumptions in these papers differ significantly. However, bid shading can also be useful in our model. Without bid shading, a truthful bidder might just win too many low-valued impressions at a high price in the early phases until his campaign target is reached. Bid shading makes the bidder lose on low-valued impressions in the early phases saving part of his campaign target for later phases of the campaign. We analyze the gains that a bidder has from unilateral manipulation and use this to characterize environments that are more or less robust against strategic manipulation in an online second-price auction.

Result 4. *If a bidder shades his value unilaterally while all others bid truthful, then this bidder's utility increases significantly when competing with only a small number of bidders. With an increasing number of bidders, bid shading quickly leads to utility losses as bidders become losing when shading too much.*

Table 5 summarizes the median of the relative differences in utility due to bid shading. In other words, we report the ratio of the utility that a single bidder would get by shading his bid by a certain factor vs. bidding truthfully as all other bidders (i.e., 1.0 represents no difference, >1.0 a surplus, and <1.0 a deficit). The table shows the median as the empirical distribution is not symmetric but very skewed. The mean might give a wrong impression about the benefits of shading, since it is very sensitive to extreme values.

The results in Table 5 reveal that with only a small number of bidders competing for an impression a bidder can gain by unilaterally deviating from truthful bidding. With higher competition

Table 5. Median Utility Gain by Unilaterally Deviating from Truthful Bidding

Number of Bidders	Bid Shading Factor 0.9	0.8	0.7	0.6	0.5	0.25
3	1.00	1.19	1.41	1.63	1.76	1.90
4	1.00	1.07	1.17	1.16	1.07	0.57
5	1.00	1.00	1.00	0.75	0.29	0.00
6	1.00	1.00	0.86	0.32	0.00	0.00
8	1.00	0.96	0.16	0.00	0.00	0.00
10	1.00	0.76	0.00	0.00	0.00	0.00
12	1.00	0.61	0.00	0.00	0.00	0.00
15	1.00	0.00	0.00	0.00	0.00	0.00

the gains quickly become losses and a bidder has to be careful not to lose against competing bidders. Overall, this suggests that the potential gain from manipulation might not be worth the risk to lose in markets with a medium or large number of bidders competing per impression. To bid strategically, a bidder needs to understand the competition that he can expect for different types of impressions. The analysis suggests that truthful bidding is no equilibrium in markets with few bidders where bidders have a campaign target. How much they should shade in an online market in equilibrium is a different question and one that is much harder to answer if the auction is modeled as a dynamic Bayesian game, which is standard in auction theory.

7 CONCLUSION

A wide-spread argument for the second-price rule in display ad, or keyword auctions with one slot only, is that they are incentive-compatible (Edelman and Schwarz 2010). The argument relies on the separability of valuations for different impressions. It is unrealistic, however, to assume that the value for the first impression that an advertiser buys is equivalent to the marginal value of an impression after having bought already several thousand impressions for a campaign. Actually, advertisers typically have a campaign target or a budget determining how much they want to spend (Berg et al. 2010). They either cannot or do not want to spend more money beyond this target.

In this article, we analyze the allocative efficiency and bidders' utility of online display ad auctions when bidders have a campaign target and report truthfully. We introduce a simple value function that reflects the campaign target, but still allows for a standard quasi-linear (i.e., payoff-maximizing) utility function. When the impressions all arrive at the same time, they could be allocated optimally using a mathematical program.

The online display ad auction can be seen as a dynamic algorithm to solve this computationally hard allocation problem. We compare the online to the optimal offline solution, assuming that bidders report their valuations truthfully at each point in time. While the offline VCG mechanism would be truthful, the online mechanism is not. First, we show that the worst-case efficiency of the online allocation in our model is 1/2 of the optimal welfare. Then, we provide results of numerical experiments showing that the average efficiency of the auction is surprisingly high on average and increases with the number of bidders.

To compare the revenue and the utility of the bidders in the online and offline setting, we compute core-payments from the offline allocation problem as a baseline. They describe the minimal competitive equilibrium prices such that no coalition of losing bidders could outbid the winners and provide a useful baseline in a model with payoff-maximizing bidders. We prove that these core payments can be higher than the VCG payments in our model. In the simulations, we find that with low competition per impression, the bidders pay substantially more in the online auction as compared to the offline auction. With an increasing number of bidders, however, the prices converge to the valuations and the differences between online display ad auctions and the hypothetical offline world are small.

Finally, we ask the question how much bidders could gain from bid shading given all others report truthfully. This analysis helps understand the robustness of the second-price rule in online markets. In line with the earlier results, the gains from such manipulations are high in markets with only few competitors per impression. In contrast, in markets with many bidders an unilateral bid shading can make a bidder lose who would otherwise win resulting in lower utility compared to a truthful strategy.

Overall, the analysis indicates that in markets with many competing advertisers per impression, manipulation is risky and there are little incentives for deviations from simple truthful bidding. There are, however, markets and types of impressions where few bidders compete. In these markets advertisers have incentives to manipulate in an attempt to increase overall utility in a model where the preferences for individual impressions are not separable as has been assumed in past literature.

As a note of caution, our simulation model differs from real-world display ad auctions where the arrival rates of different impression types vary substantially over the day and there are many more impression types. Our simulation model removes this noise from the daily fluctuations. While the absolute numbers on efficiency losses do not have external validity for this reason, the model allows us to better understand causal relationships between the bidding strategy and the efficiency and revenue of the auctions.

APPENDICES

A PROOFS

Proof of Proposition 1:

To prove this, we rely on the results by Milgrom and Strulovici (2009), who show that the gross substitutes condition (GS) guarantees VCG to be in the core, and of Fujishige and Yang (2003), who prove the equivalency of $M^\natural$-concavity to GS for $\{0,1\}^n$ valuation functions.

Definition 1 (Shioura and Tamura 2015). A function $f : \mathbb{Z}^n \to \mathbb{R} \cup \{-\infty\}$ is said to be $M^\natural$-concave if it satisfies the following exchange property: $\forall x, y \in dom f, \forall i \in supp^+(x-y), \exists j \in supp^-(x-y) \cup \{0\}$:

$$f(x) + f(y) \leq f(x - \chi_i + \chi_j) + f(y + \chi_i - \chi_j),$$

where $N = \{1, \dots, n\}$, dom $f = \{x \in \mathbb{Z}^n | f(x) > -\infty\}$, $supp^+(x) = \{i \in N | x_i > 0\}$, $supp^-(x) = \{i \in N | x_i < 0\}$ for $x \in \mathbb{Z}^n$, χ_i is the characteristic vector of $i \in N$, and $\chi_0 = \mathbf{0}$.

While a purely additive (or linear) valuation functions is always $M^\natural$-concave and therefore always GS (Shioura and Tamura 2015), our valuation function v with a campaign target B_i is not $M^\natural$-*concave*.

Proposition 3. *The valuation function v is not $M^\natural$-concave. As a result GS is violated and the VCG payments do not necessarily lie in the core.*

Proof. We prove the proposition by providing a counterexample for the $M^\natural$-concavity of the value function v. Consider $n > 6$ impressions $K = \{1, \dots, n\}$ where the first impression is of type t_1, impressions $2, \dots, n-2$ are of type t_2 and impressions $n-1, n$ are of type t_3. Let $b \in I$ be a bidder with a campaign target of $B_b = n$ and valuations $v_{bt_1} = 2, v_{bt_2} = 1, v_{bt_3} = \frac{n-1}{2}$. Let $x = (1, 1, \ldots, 1, 0, 0)$ and $y = (0, 0, \ldots, 0, 1, 1)$ be two allocations. We show that for $i = 1 \in supp^+(x - y)$ there is no $j \in supp^-(x - y) \cup \{0\}$ for which the exchange property holds.

For $j = 0$, we get

$$v_b(x) + v_b(y) = n - 1 + n - 1 = 2n - 2 \nleq 2n - 3 = n - 3 + n = v_b(x - \chi_1) + v_b(y + \chi_1), \tag{4}$$

and for $j \in \{n - 1, n\}$, we get

$$\begin{aligned} v_b(x) + v_b(y) = n - 1 + n - 1 = 2n - 2 \nleq \frac{3}{2}(n + 1) = n + \frac{n-1}{2} + 2 \\ = v_b(x - \chi_1 + \chi_j) + v_b(y + \chi_1 - \chi_j). \end{aligned} \tag{5}$$

Hence, the valuation function is not $M^\natural$-concave. □

Proof of Proposition 2:

Proof. We first show that the social welfare of an online allocation is bounded from below by one half of the welfare of an optimal offline solution. Let X be an optimal offline allocation and Y be an allocation in the online scenario. Consider an impression k for which there exists a bidder i that is assigned k in X such that this assignment contributes v^* to the total social welfare in X, but for which this impression only contributes $v' < v^*$ to the total welfare in Y. Since i is assigned the impression in the offline allocation, $v_{it(k)} \geq v^*$ has to hold. Consider all remaining impressions $K' = K \setminus \{k\}$. Bidder i obtains a combined value of at least $B - v'$ by being assigned impressions from K' in Y, since otherwise i would place a bid greater than v' for k (as he has higher value for this impression). However, he only obtains a combined value for impressions K' of at most $B - v^*$ in X. Let V^Y be the sum of values of all impressions that bidder i obtains in Y and let V^X be the sum of values of all impressions for bidder i in X. From above it follows that $V^X - V^Y \leq v' - v^*$ or $v^* \leq v' + V^Y - V^X$.

Let $K^Y \subseteq K'$ be the set of impressions obtained by bidder i in Y that he does not obtain in X and and let $K^X \subseteq K'$ be the set of impressions obtained by i in X but not in Y. In the worst case, all bidders who obtained the impressions in K^Y in X (i.e., those that bidder i does not get there) have no value for any other impressions and hence are not assigned any impressions in Y. Thus, in this case, none of these values contribute to the total welfare in the online scenario. Additionally, there may be no other bidder who values impression in K^X, such that these are unassigned in Y as well and hence also non of these values contribute to the total welfare in Y.

Combining these results, in the worst case, the total welfare contributed by assigning impressions of $K^X \cup K^Y \cup \{k\}$ in Y as compared to X is at least

$$\frac{0 + V^Y + v'}{V^X + V^Y + v^*} \geq \frac{V^Y + v'}{2V^Y + v'} \geq \frac{1}{2}, \tag{6}$$

i.e., at least half of the social welfare in X. Continuing iteratively with all other impressions and bidders yields the lower bound.

In the following, we show that this bound is tight by proving an upper bound of one half, as well. Consider the following example with two bidders, 1, 2 with identical campaign targets $B = B_1 = B_2$, two impressions k, l and valuations $v_{1t(k)} = v_{1t(l)} = v_{2t(k)} = B, v_{2t(l)} = 0$. Then, in an optimal allocation, bidder 1 obtains impression l while bidder 2 obtains k for a combined welfare of $2B$. However, if impression k is auctioned off first to bidder 1, then the combined welfare is only B. □

B TABLES

Table 6. Second Setup: Linear Regression of Relative Utility w.r.t. Grouped Number of Bidders, Considering Low Competition (3 or 4 Bidders) as Baseline, and Impression Types

	Relative Utility
(Intercept)	0.56^{***}
Medium Competition (5 or 6 Bidders)	0.09^{***}
	(0.01)
	(0.00)
High Competition (7–15 Bidders)	0.09^{***}
	(0.01)
Number of Impression Types	-0.02^{***}
	(0.00)
Medium Competition (5 or 6 Bidders): Number of Impression Types	0.03^{***}
	(0.00)
High Competition (7–15 Bidders): Number of Impression Types	0.04^{***}
	(0.00)
R^2	0.53
Adj. R^2	0.52
Num. obs.	29,250
RMSE	0.13

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Table 7. Total Online Utility

Number	Number of Impression Types								
of Bidders	2	3	4	5	6	7	8	9	10
3	40,734	45,062	44,642	45,530	46,673	48,395	48,199	49,116	46,988
4	37,571	40,962	40,335	38,229	44,560	46,320	44,103	43,193	47,095
5	28,483	34,428	35,865	41,202	37,309	42,439	39,951	38,390	38,780
6	25,921	35,127	38,125	36,312	38,366	40,366	38,119	41,001	41,469
7	28,234	30,875	33,798	37,257	37,641	38,739	37,181	38,742	38,385
8	26,675	30,134	33,880	35,829	32,693	33,778	37,702	38,515	38,493
9	26,408	29,877	34,357	34,507	34,086	38,307	37,533	40,122	37,956
10	25,776	29,531	31,650	30,504	33,351	36,128	36,047	36,061	38,566
11	25,184	29,040	31,092	30,063	36,010	35,641	33,715	34,096	39,198
12	25,516	31,997	29,478	33,565	29,909	35,281	36,616	37,249	38,755
13	25,136	26,383	31,952	32,635	31,990	33,448	39,255	38,439	37,859
14	25,680	28,021	30,192	28,729	31,376	33,692	34,767	37,305	34,658
15	24,327	30,186	27,907	31,697	34,706	36,157	35,277	35,423	38,094

Table 8. Total Offline Utility

Number of Bidders	Number of Impression Types 2	3	4	5	6	7	8	9	10
3	90,784	103,177	105,214	112,082	112,177	117,648	117,031	118,908	117,629
4	72,089	76,174	85,044	81,982	97,694	101,528	112,197	104,977	122,863
5	46,118	53,671	54,946	64,117	55,744	64,959	59,420	56,543	56,633
6	39,892	52,117	55,291	49,613	53,688	54,367	50,307	54,477	52,646
7	45,370	45,884	46,258	48,307	47,460	49,283	46,107	49,072	46,980
8	46,317	44,675	44,334	44,909	41,045	41,736	48,437	46,027	45,924
9	39,710	40,575	43,518	42,772	41,419	46,169	45,096	47,066	44,026
10	39,848	41,986	41,360	38,702	39,832	44,086	42,197	41,878	45,021
11	43,348	42,780	39,470	38,688	42,631	41,395	38,891	39,088	43,985
12	41,513	44,886	38,585	41,169	36,530	41,763	42,311	42,006	43,517
13	40,487	36,363	40,947	40,679	38,498	39,310	45,505	43,987	41,688
14	38,152	38,977	38,338	34,980	37,206	39,643	39,748	42,210	37,604
15	36,867	41,946	34,861	36,738	40,781	42,016	39,847	38,974	41,870

Table 9. Minimum Observed Efficiency

Number of Bidders	Arrival Rate	Number of Impression Types 2	3	4	5	6	7	8	9	10
3	clustered	0.84	0.78	0.92	0.80	0.83	0.85	0.86	0.89	0.91
3	uniform	0.90	0.88	0.90	0.90	0.89	0.91	0.90	0.92	0.89
4	clustered	0.86	0.84	0.89	0.86	0.87	0.87	0.86	0.88	0.89
4	uniform	0.91	0.93	0.91	0.88	0.91	0.89	0.88	0.88	0.92
5	clustered	0.87	0.90	0.89	0.90	0.89	0.88	0.90	0.90	0.85
5	uniform	0.90	0.91	0.91	0.90	0.93	0.92	0.92	0.90	0.89
6	clustered	0.93	0.87	0.89	0.89	0.91	0.90	0.90	0.91	0.92
6	uniform	0.94	0.92	0.94	0.91	0.92	0.88	0.94	0.91	0.94
8	clustered	0.90	0.91	0.94	0.92	0.91	0.90	0.90	0.92	0.93
8	uniform	0.94	0.94	0.93	0.94	0.94	0.94	0.92	0.93	0.95
10	clustered	0.92	0.93	0.93	0.91	0.92	0.93	0.94	0.96	0.93
10	uniform	0.91	0.94	0.96	0.93	0.92	0.95	0.94	0.95	0.96
12	clustered	0.96	0.92	0.89	0.96	0.92	0.96	0.94	0.95	0.95
12	uniform	0.97	0.96	0.95	0.95	0.93	0.96	0.94	0.95	0.95
15	clustered	0.94	0.95	0.94	0.89	0.93	0.91	0.94	0.95	0.96
15	uniform	0.96	0.96	0.97	0.96	0.94	0.94	0.95	0.96	0.96

Table 10. Average Observed Efficiency

Number of Bidders	Arrival Rate	Number of Impression Types 2	3	4	5	6	7	8	9	10
3	clustered	0.98	0.98	0.99	0.98	0.99	0.99	0.99	0.99	0.99
3	uniform	0.98	0.97	0.98	0.98	0.98	0.98	0.99	0.99	0.99
4	clustered	0.98	0.97	0.98	0.97	0.97	0.97	0.97	0.97	0.97
4	uniform	0.98	0.97	0.97	0.97	0.96	0.96	0.96	0.96	0.96
5	clustered	0.98	0.99	0.98	0.97	0.97	0.97	0.97	0.98	0.97
5	uniform	0.98	0.98	0.97	0.97	0.97	0.96	0.97	0.97	0.97
6	clustered	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
6	uniform	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.98
8	clustered	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.99
8	uniform	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98	0.99
10	clustered	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
10	uniform	0.99	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99
12	clustered	1.00	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99
12	uniform	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
15	clustered	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
15	uniform	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99

REFERENCES

Santiago R. Balseiro, Omar Besbes, and Gabriel Y. Weintraub. 2015. Repeated auctions with budgets in ad exchanges: Approximations and design. *Manage. Sci.* 61, 4 (2015), 864–884.

Jordan Berg, Amy Greenwald, Victor Naroditskiy, and Eric Sodomka. 2010. A knapsack-based approach to bidding in ad auctions. In *Proceedings of the European Conference on Artificial Intelligence (ECAI'10)*, vol. 215. 1013–1014.

Deeparnab Chakrabarty, Yunhong Zhou, and Rajan Lukose. 2008. Online knapsack problems. In *Proceedings of the Workshop on Internet and Network Economics (WINE'08)*.

Jianqing Chen and Jan Stallaert. 2014. An economic analysis of online advertising using behavioral targeting. *MIS Quart.* 38, 2 (2014), 429–449.

R. Day and S. Raghavan. 2007. Fair payments for efficient allocations in public sector combinatorial auctions. *Manage. Sci.* 53, 9 (2007), 1389–1406.

Liu De, Jianqing Chen, and Andrew B. Whinston. 2010. Ex ante information and the design of keyword auctions. *Info. Syst. Res.* 21, 1 (2010), 133–153.

Shahar Dobzinski, Ron Lavi, and Noam Nisan. 2008. Multi-unit auctions with budget limits. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS'08)*. IEEE, 260–269.

Benjamin Edelman and Michael Schwarz. 2010. Optimal auction design and equilibrium selection in sponsored search auctions. *Amer. Econ. Rev.* 100, 2 (2010), 597–602.

Satoru Fujishige and Zaifu Yang. 2003. A note on Kelso and Crawford's gross substitutes condition. *Math. Oper. Res.* 28, 3 (2003), 463–469.

Vijay Krishna. 2009. *Auction Theory*. Academic Press.

Alberto Marchetti-Spaccamela and Carlo Vercellis. 1995. Stochastic on-line knapsack problems. *Math. Program.* 68, 1–3 (1995), 73–104.

Andreu Mas-Colell, Michael Dennis Whinston, Jerry R. Green et al. 1995. *Microeconomic Theory*, vol. 1. Oxford University Press, New York.

Paul Milgrom and Bruno Strulovici. 2009. Substitute goods, auctions, and equilibrium. *J. Econ. Theory* 144, 1 (2009), 212–247.

A. Shioura and A. Tamura. 2015. Gross substitutes condition and discrete concavity for multi-unit valuations: A survey. *J. Operat. Res. Soc. Japan* 58, 1 (2015), 61–103.

William Vickrey. 1961. Counterspeculation, auctions, and competitive sealed tenders. *J. Finance* 16, 1 (1961), 8–37.

Sha Yang and Anindya Ghose. 2010. Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Market. Sci.* 29, 4 (2010), 602–623.

Shuai Yuan, Jun Wang, and Xiaoxue Zhao. 2013. Real-time bidding for online advertising. In *Proceedings of the 7th International Workshop on Data Mining for Online Advertising (ADKDD'13)*, Esin Saka, Dou Shen, Bin Gao, Jun Yan, and Ying Li (Eds.). ACM Press, New York, New York, 1–8. DOI: https://doi.org/10.1145/2501040.2501980

Zenith. 2017. Internet advertising expenditure to exceed US $200bn this year. Retrieved from https://www.zenithmedia.com/internet-advertising-expenditure-to-exceed-us200bn-this-year/.

Weinan Zhang, Shuai Yuan, and Jun Wang. 2014. Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1077–1086.

Xia Zhao and Ling Xue. 2012. Competitive target advertising and consumer data sharing. *J. Manage. Info. Syst.* 29, 3 (2012), 189–222.

Wen-Yuan Zhu, Wen-Yueh Shih, Ying-Hsuan Lee, Wen-Chih Peng, and Jiun-Long Huang. 2017. A gamma-based regression for winning price estimation in real-time bidding advertising. In *Proceedings of the IEEE International Conference on Big Data (BigData'17)*. IEEE, 1610–1619. DOI: https://doi.org/10.1109/BigData.2017.8258095

Received June 2018; revised March 2019; accepted April 2019

4 Market Design for Renewable Energy Auctions

Peer-Reviewed Journal Paper
Title: Market Design for Renewable Energy Auctions: An Analysis of Alternative Auction Formats
Authors: M. Bichler, V. Grimm, S. Kretschmer, P. Sutterer
In: Energy Economics
Abstract: Auctions are widely used to determine the remuneration for renewable energies. They typically induce a high concentration of renewable energy plants at very productive sites far-off the main load centres, leading to an inefficient allocation as transmission line capacities are restricted but not considered in the allocation, resulting in an inefficient system configuration in the long run. To counteract these tendencies effectively, we propose a combinatorial auction design that allows to implement regional target capacities, provides a simple pricing rule and maintains a high level of competition between bidders by permitting package bids. By means of extensive numerical experiments we evaluate the combinatorial auction as compared to three further renewable energy source auction designs: the current German nationwide auction design, a simple nationwide auction, and regional auctions. We find that if bidders benefit from high enough economies of scale, the combinatorial auction design implements system-optimal target capacities without increasing the average remuneration per kWh as compared to the current German auction design. The prices resulting from the combinatorial auction are linear and anonymous for each region whenever possible, while minimal personalised markups on the linear prices are applied only when necessary. We show that realistic problem sizes can be solved in seconds, even though the problem is computationally hard.

Energy Economics 92 (2020) 104904

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

Market design for renewable energy auctions: An analysis of alternative auction formats

Martin Bichler [a], Veronika Grimm [b,c,*], Sandra Kretschmer [b,c], Paul Sutterer [a]

[a] *Department of Informatics, Technical University of Munich, Boltzmannstr. 3, 85748 Garching, Germany*
[b] *Department of Business and Economics, Friedrich-Alexander University Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany*
[c] *Energy Campus Nuremberg, Fuerther Strasse 250, 90429 Nuremberg, Germany*

ARTICLE INFO

Article history:
Received 9 July 2019
Received in revised form 27 July 2020
Accepted 7 August 2020
Available online 15 August 2020

Keywords:
Renewable energy auction
Combinatorial auction
Market design

ABSTRACT

Auctions are widely used to determine the remuneration for renewable energies. They typically induce a high concentration of renewable energy plants at very productive sites far-off the main load centres, leading to an inefficient allocation as transmission line capacities are restricted but not considered in the allocation, resulting in an inefficient system configuration in the long run. To counteract these tendencies effectively, we propose a combinatorial auction design that allows to implement regional target capacities, provides a simple pricing rule and maintains a high level of competition between bidders by permitting package bids. By means of extensive numerical experiments we evaluate the combinatorial auction as compared to three further RES auction designs, the current German nationwide auction design, a simple nationwide auction, and regional auctions. We find that if bidders benefit from high enough economies of scale, the combinatorial auction design implements system-optimal target capacities without increasing the average remuneration per kWh as compared to the current German auction design. The prices resulting from the combinatorial auction are linear and anonymous for each region whenever possible, while minimal personalised markups on the linear prices are applied only when necessary. We show that realistic problem sizes can be solved in seconds, even though the problem is computationally hard.

1. Introduction

Auctions are widely used in Europe and even worldwide to determine the remuneration for renewable energy sources (RES). Numerous countries have implemented auction systems to determine feed-in tariffs (FITs) and -premiums (FIPs), deductions or other subsidies for electricity from renewable power plants, among them Brazil, China, Denmark, France, Portugal, Germany, South Africa and the United Kingdom, to name but a few.

Compared to fixed FITs and FIPs, auctions – if designed and implemented appropriately – have the potential to reduce remuneration and thus avoid overcompensation (de Vos and Klessmann, 2014; del Río and Linares, 2014; Mora et al., 2017). Renewable energy auctions are often considered economically efficient, since the resulting remuneration is competitively determined, close to the bidders' true cost and capacity expansion can be steered more effectively (Cozzi, 2012; del Río and Linares, 2014; Held et al., 2014; Kreiss et al., 2017; Mora et al., 2017). However, this typically induces a high concentration of renewable energy power plants at the most productive (i.e. windy or sunny) sites, which are often located far from the main load centres (Gerlach et al., 2015; Ibrahim et al., 2011; IRENA and CEM, 2015).

To counteract these tendencies, some countries, e.g. Germany or Uruguay, have established mechanisms that modulate support levels according to the location of a RES plant to induce a broader regional distribution of plants, in particular closer to main load centres (IRENA and CEM, 2015; Federal Ministry for Economic Affairs and Energy, 2016). Yet such location-specific auction mechanisms frequently fail to account for relevant aspects like potential network congestion arising from renewables expansion. Overall, the German mechanism, the reference yield model (Referenzertragsmodell, REM), has been criticised to reduce incentives to build wind power plants at efficient sites, while it also does not provide an effective means to steer generation capacity to particular target regions: the capacities awarded under the current German wind auction design with the REM are far from the envisaged targets in the individual regions (Güsewell, 2016; Jürgens, 2017).

In this paper, we build on advances in auction design to assess the potential of combinatorial auctions to induce efficient locational choices for RES and thereby reach the aforementioned efficiency goals. Combinatorial auction designs have already been successfully applied in

* Corresponding author at: Department of Business and Economics, Friedrich-Alexander University Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany.
E-mail addresses: bichler@in.tum.de (M. Bichler), veronika.grimm@fau.de (V. Grimm), sandra.kretschmer@fau.de (S. Kretschmer), paul.sutterer@tum.de (P. Sutterer).

https://doi.org/10.1016/j.eneco.2020.104904
0140-9883/© 2020 Published by Elsevier B.V.

transportation, logistics and for spectrum sales (Bichler and Goeree, 2017). In particular, combinatorial auctions allow to determine target capacities for each region and, at the same time, maintain a sufficient level of competition between bidders. Moreover, bidders can express their economies of scale across projects in one or more regions via package bids, i.e. they can submit package bids for any arbitrary combination of their planned projects. The winning bids are then assigned such that the most efficient combinations of project portfolios are selected while the tendered capacity per region is met and not overly exceeded.

On the downside, package bids typically do not allow for linear and anonymous prices. Non-linear and personalised package prices, used e.g. in spectrum auctions, are often perceived as unfair (Bichler and Goeree, 2017). For our analyses, we therefore leverage the fact that RES auctions are large and the resulting non-convexities in the associated optimisation problems small. Our proposed auction design computes linear and anonymous prices whenever possible, and minimal mark-ups for the winning bidders on top of these anonymous linear prices when exact linear and anonymous prices are impossible. I.e. we determine one price per region for all bidders (potentially with a minimal mark-up for some winning bidders to avoid losses). The overall design draws on theoretical insights from the literature on pricing in combinatorial auctions (Adomavicius et al., 2012; Bichler et al., 2017, 2010; Guo et al., 2012). Even though the allocation pricing is computationally hard, we can solve realistic problem sizes in minutes, as we will show.

With this in mind, the focus of this paper is to assess and evaluate the impact of different RES auction designs on the resulting allocative efficiency and subsidy payments, i.e. the cost for the taxpayer. The short time span since most RES auction schemes are in place and the limited data availability renders an empirical evaluation of different RES auction designs infeasible. Moreover, since combinatorial auctions have not been used for this purpose in practice, a counterfactual analysis of different auction designs with field data is impossible. We therefore conduct extensive numerical experiments, with our analyses based on the case of RES auctions in Germany.

Germany constitutes an excellent starting point for our study, as both comprehensive market data as well as information on system-optimal RES locations are available to calibrate the numerical model. Following the European Commission's guidelines to use auction schemes for RES support (European Commission, 2014), Germany's renewable energies support scheme was revised in 2017. Since then, the capacities and the remuneration of all renewable energy plants exceeding 750 kilowatt (kW) capacity are determined via auctions (see section 22 Renewable Energy Sources Act, EEG, 2017). Pre-defined capacities are auctioned off to the lowest remuneration rates asked for in up to four auctions per technology a year (see section 32 EEG, 2017). Like in most countries with RES auctions, winning bidders are awarded a sliding FIP per kilowatt-hour (kWh), i.e. a subsidy covering the difference between the current electricity market price and the awarded bid price, if the market price lies below the awarded bid price (Wigand et al., 2016).

In our numerical study, we focus on onshore wind auctions, as onshore wind power is the capacity-wise largest renewable energy technology in Germany (section 28 EEG, 2017; Federal Ministry for Economic Affairs and Energy, 2018). We compare the current design, a nationwide auction with REM, to (i) a simple nationwide auction design, (ii) a regional auction design that implements the desired regional capacities and (iii) a combinatorial auction design that implements the desired regional target capacities but maintains a sufficient level of competition. The regional target capacities are taken from a study by Grimm et al. (2017), who determine the optimal RES locations in Germany accounting for the location of load centres and available network capacities.

We find that for reasonable synergies among projects, the proposed combinatorial auction design implements the system-optimal regional target capacities without considerably increasing the resulting remuneration compared to the current German auction design. This is surprising, given that the combinatorial auction exactly implements regional target capacities at less productive sites closer to demand centres, while the current nationwide auction design does not face those constraints. Notably, this remuneration does not even include long-run cost savings resulting from lower redispatch and network investment requirements when regional target capacities are met, the cost for which easily exceed a billion Euro per year (e.g. €1.5bn in 2017) and are largely caused by the regional mismatch of supply and demand (Federal Network Agency and Federal Cartel Office, 2019). Grimm et al. (2017) show that a system-optimal allocation of RES in Germany enables efficiency gains of up to €2.6bn p.a. Although we focus on a particular application to the current German wind auctions, key insights can be obtained for RES auctions in general.

Using auctions as a support mechanism for the deployment of renewable energies is no new phenomenon in Europe or even worldwide. Numerous European countries have implemented renewable energy auction systems even before 2017,[1] among them e.g. Denmark, France, Germany, Ireland, Italy, the Netherlands, Portugal and the United Kingdom – with mixed results.[2]

Across these eight countries, several trends were apparent: all countries required bids to be submitted in capacity (kW or MW) and mainly conducted technology-specific auctions (for this and the following refer to Wigand et al., 2016). Most countries (Denmark, Germany, Italy, the Netherlands and the UK) awarded sliding FIPs, while France, Ireland and Portugal awarded fixed FITs. Equally favoured was the pay-as-bid pricing rule, which was applied by all countries but the Netherlands and the UK, who opted for a uniform pricing rule. Furthermore, most countries chose price-only auction schemes for bid evaluation, while France and Portugal additionally consider further criteria like a plant's carbon footprint and the development of industrial clusters, respectively. The duration of the support payments ranged from 12 years in Denmark to up to 30 years in Italy.

Many auction design choices depend on and reflect political priorities, the country's market situation and socio-institutional landscape. Consequently, the designs varied regarding further design elements like technology-specific vs. open auctions, volume vs. budget caps, restrictions of bid sizes, concentration rules, frequency of auctions, ceiling prices, pre-qualification criteria, realisation periods, penalties and bid bonds (for more on this see IRENA and GWEC, 2012; Wigand et al., 2016).

Lacking grid connections, inadequate bid bonds, low support payments as a result of speculative bids and missing penalty mechanisms were some of the main problems that disincentivised project realisation (del Río and Linares, 2014; Förster, 2016; Gallachóir et al., 2009; McLean et al., 2007; Tiedemann et al., 2016; Wigand et al., 2016). Up until 2017, only Denmark and Portugal had achieved realisation rates of 100%, although with serious delays in Portugal (Peña Cabra, 2014; Wigand et al., 2016). All eight countries reported lower support levels under their respective auction schemes compared to previous support mechanisms, though remuneration levels still remained comparatively high in most countries (Fitch-Roy and Woodman, 2016; Heer et al., 2007; Kitzing and Wendring, 2015; Negri, 2015; Wigand et al., 2016).

Our research builds on several strands of the literature. Building on experiences with RES auctions in various countries and on insights from auction theory, there is large consent on the importance of certain design elements for the general success of renewable energy auctions. An overview and discussion is provided in e.g. Cramton (2010), IRENA and CEM (2015), Klemperer (2004), Maurer and Barroso (2011) or del

[1] The year as of which the European Commission's State Aid Guidelines recommend all member states to apply competitive auction schemes for renewables support (European Commission, 2014)

[2] Ireland conducted RES auctions as early as 1995 until 2003, Denmark, France, the Netherlands, Italy and Germany introduced auctions for some RES already in 2004, 2011, 2013 and 2015, respectively, while in Portugal, auctions took place between 2006 and 2008 for wind and biomass and were just introduced for PV in 2019. The UK introduced its current auction scheme in 2014.

Table 1
Design elements of the German onshore wind auction design.

Auction design element	Implementation
Product	Installed capacity (MW)
Pricing rule	Pay-as-bid and uniform price sealed-bid auction (for BEG)
Type	Price-only multi-item auction
Auctioned volume	2800 MW per year, i.e. 700–1000 MW per round
Remuneration scheme	Energy-related remuneration (capacity is tendered, electricity is remunerated)
Price ceiling	7 $^{ct}/_{kWh}$ in 2017; from 2018: average of highest accepted bid in the last three rounds, increased by 8% (6.3 $^{ct}/_{kWh}$ in 2018, 6.2 $^{ct}/_{kWh}$ in 2019)
Pre-qualification requirements	Bid bond of 30 $^{€}/_{kW}$ of installed capacity (for BEG: 15 $^{€}/_{kW}$, secondary bid bond of 15 $^{€}/_{kW}$ upon winning) BImSchG-approval 3 weeks before auction
Frequency	3 to 4 auctions a year (every 2–4 months)
Concentration rules	Min. 750 kW Max. 6 bids for max. 18 MW in total for BEG
Penalties	10 $^{€}/_{kW}$ after 24 (48) 20 $^{€}/_{kW}$ after 26 (50) months of delay (for BEG) 30 $^{€}/_{kW}$ after 28 (52)
Form of support	Sliding FIP per kWh
Support duration	20 years

Source: Own elaboration.

Río et al. (2015). In particular, several studies stress the need for enforceable penalties to preclude project non-realisation, and advise in this regard to additionally include financial securities, fixed construction deadlines as well as prequalification criteria that require projects to be in an advanced planning stage, i.e. to conduct 'late auctions' (Anaya and Pollitt, 2015; de Jager and Rathmann, 2008; del Río and Linares, 2014; Maurer and Barroso, 2011; Mora et al., 2017; Toke, 2015). Several contributions point to a significant trade-off between cost-efficient support levels, reaching capacity expansion targets and actor diversity (del Río, 2017; Grashof, 2013; Hauser et al., 2014; Hauser and Kochems, 2014).

The optimal allocation of RES capacity remains largely untouched in the literature on RES auctions. While traditionally, under fixed feed-in tariffs, locational choice for renewable energy plants only depended on site specific weather conditions, several recent studies, among them Benz et al. (2015) and Grimm et al. (2017), illustrate the systemic optimality of a decentralised allocation of generation capacity that additionally accounts for existing network infrastructure and potentially arising network constraints. They show that closer proximity to main demand centres can significantly reduce prospective network congestion and, ultimately, the need for transmission line expansion. Ackermann et al. (2001) and Amado et al. (2017) support this notion and show that especially renewable energies are very well suited for distributed generation and smart grids.

Current studies using numerical experiments and simulations to compare or devise auction designs in an energy-related context have so far mainly focused on day-ahead auctions (e.g. Contreras et al., 2001; Fernandez-Blanco et al., 2014; Kardakos et al., 2013), local reserve energy markets (e.g. Rosen and Madlener, 2013), inter-grid power auctions (e.g. Zhang et al., 2014), carbon allowance auctions (e.g. Tang et al., 2017), energy contract auctions (e.g. Barroso et al., 2011) or an agent-based comparison of pay-as-bid vs. uniform price wind auctions (Anatolitis and Welisch, 2017). To our knowledge, we are the first to combine the findings on RES auction design and a system-optimal distribution of generation capacity to analyse whether minor adjustments to an existing auction design can lead to an improved regional distribution of generation capacities and compare several auction mechanisms with regard to the resulting FIPs.

The remainder of the paper is structured as follows: Section 2 illustrates the current situation and wind auction design in Germany, before we describe our model and the analysed auction designs in Section 3. Section 4 provides information on the data and experimental design, while we present our results in Section 5. Section 6 concludes.

2. Onshore wind auctions: A discussion based on the German case

As outlined above, the aim of our study is to conduct a counterfactual analysis of different auction designs to assess the effects of implementing system-optimal RES locations on the resulting level of remuneration.

To make the effects quantifiable, we conduct the analysis using data from the German onshore wind auctions. Germany is an ideal example to illustrate the comparative performance of different auction designs, as not only comprehensive market data is available for calibration, but also information on system-optimal RES locations (see Section 4.1). We therefore start from the current German wind auction design, before we discuss and evaluate alternative auction designs. Although we focus on a particular application to the current German wind auctions, key insights can be obtained for RES auctions in general.

As a starting point to motivate and outline the structural background of our model, this section provides a brief description of the legal framework for wind auctions in Germany, the currently implemented reference yield model, insights on the system-optimal allocation of RES in Germany as well as a discussion of alternative RES auction designs. Building on this, we specify our model and report our results in the subsequent sections.

2.1. Legal framework and results

In Germany, the EEG, 2014 introduced RES auctions for the first time. Since 2017, the annual auction volume for onshore wind energy has been 2.8 Gigawatt (see section 28 EEG, 2017). It is divided into four auctions, which take place quarterly. The rules for each single tender auction are as follows: (1) For each tender, the Federal Network Agency defines an exact auction volume of onshore wind capacity to be installed; (2) subsequently, bidders compete on the remuneration per kWh of fed-in electricity; (3) the bidders offering the lowest remuneration per kWh win, until full capacity is reached; remuneration is guaranteed for 20 years as follows: (4.a) winners receive their offered remuneration per kWh or (4.b) receive the highest accepted remuneration per kWh if they are a so-called 'citizen energy cooperative'.

Citizen energy cooperatives (Buergerenergiegesellschaften, BEG) are local communities that intend to build a wind power plant (Deutsche WindGuard, 2017). By allowing BEG to participate in the auctions, the legislator wants to foster bidder diversity and thus increase local acceptance and regional value added. In order for them to be competitive, BEG are granted some simplifications in the auction procedure. Most importantly, while institutional bidders face a pay-as-bid auction, BEG are subject to a uniform price auction and receive the tariff of the highest accepted bid in the same round (sections 3(51) and 36 g (5) EEG, 2017).[3]

The final FIP paid to each winner then results from the awarded remuneration per kWh minus the average monthly electricity market price, which is called a "sliding FIP" (annex to section 23a EEG). Since the final FIPs and thus the subsidy payments by the regulator depend on the hourly electricity market price, which we do not simulate in our model, we focus our analysis and discussion on the offered and awarded remuneration per kWh, and not the resulting FIPs. Note that this does not limit the interpretation of our results: the difference between the remuneration levels resulting from the various auction designs corresponds exactly to the difference in subsidy payments by the regulator.

[3] In addition, BEG also face a lower required collateral and penalty fees and are allowed a longer realisation period (see sections 36 and 36 g EEG).

In 2017, a price ceiling of 7 $^{ct}/_{kWh}$ was set, above which bids were not accepted. This price ceiling was lowered to 6.3 $^{ct}/_{kWh}$ in the 2018 and to 6.2 $^{ct}/_{kWh}$ for the 2019 auctions (see sections 36b and 85(1) EEG and Federal Network Agency, 2019a). Except for the BEG, who can submit at most six bids for no more than 18 MW in sum, there are no restrictions on maximum awarded capacity or number of bids. To ensure a high realisation rate, bidders must submit approval according to the Federal Immission Control Act (BImSchG) three weeks prior to the auction date (sections 36 and 104(8) EEG),[4] BImSchG (2017) as well as a bid bond of 30 $^{€}/_{kW}$ installed capacity (15 $^{€}/_{kW}$ for BEG, with another 15 $^{€}/_{kW}$ due upon winning). Failure to commission a plant within the prescribed deadline of 24 months (48 months for BEG) results in a penalty fee of 10 $^{€}/_{kW}$ after 24 (48), 20 $^{€}/_{kW}$ after 26 (50) and 30 $^{€}/_{kW}$ after 28 (52) months. After a delay of 30 (54) months or a default on the security payments, an awarded tender is withdrawn. Table 1 provides a condensed overview of the main design elements of the German onshore wind auctions.

ALGORITHM 1: Allocation and pricing rule for *National, National REM* and *Regional* auction designs

Data: Submitted bids, represented by tuples of ask price b_j and capacity y_j for a project j
Result: Set of winning bids W, allocated remuneration p_j
Determine tendered capacity D;
Sort submitted bids in ascending order by ask price b_j;
In case of a tie: sort in ascending order by capacity y_j;
while $D > 0$ **do**
    assign j to W;
    $p := b_j$;
    if *project not a BEG* **then**
        $p_j := b_j$;
    end
    $D := D - y_j$;
end
To all projects in W without price, assign $p_j = p$;

2.2. The reference yield model

Another important element of the German onshore wind auctions is the reference yield model ('Referenzertragsmodell, REM) (see annex 2 EEG). The motivation for its introduction was to create a level playing field for wind plant projects across various sites in Germany. This is supposed to foster a more even distribution of wind power plants across Germany to relieve the already heavily loaded transmission lines (Federal Ministry for Economic Affairs and Energy, 2016).

To provide incentives for plant operators to also build wind power plants in less windy areas, and not only in the particularly windy north, a lower site quality of potential plant locations is compensated by a higher remuneration: The expected electricity yield of a wind power plant at its planned location is put into proportion with the so-called reference yield, i.e. the hypothetical electricity yield this particular wind power plant would generate at a pre-defined reference site, thereby arriving at a relative site quality factor. At an 80% site, for instance, the expected yield is 20% less than at the reference site, and at a 120% site, it is 20% higher. For many current wind power plant types, the reference yield is provided by Fördergesellschaft Windenergie und andere Dezentrale Energien (FGW), 2017.

Having this in mind, bidders have to submit the bids for their onshore wind projects as if they were to be built at the reference site, so as to make bids for drastically different locations comparable regarding the wind conditions and site quality. The FIP that winning bidders then receive is their ask price adjusted by a correction factor based on the aforementioned relative site quality factor, i.e. for a less windy site the remuneration is adjusted upwards, while it is adjusted downwards for a windier one. The correction factor for various site quality levels is shown in Table 2.

2.3. System-optimal RES capacity allocation

RES support mechanisms often do not accurately account for important factors that determine the optimal spatial allocation of RES generation capacities, since the minimisation of subsidies is typically the main target criterion and indicator. While the minimisation of subsidies implies that renewable energy plants should be built at the most productive locations, we follow Grimm et al. (2019; 2018; 2016; 2017) and expand on the definition of allocative efficiency by additionally considering existing load centres as well as potential congestion of the network infrastructure that would, in the medium term, induce costs for network expansion in case of regionally concentrated RES locations. This concept naturally implies a closer proximity of generation capacity to load centres. We call the resulting allocation of generation capacity "system-optimal". Grimm et al. (2017) determine this system-optimal RES allocation for Germany using a comprehensive energy market

Table 2
Site quality and correction factors for German onshore wind auctions.

Site quality factor	≤ 70%	80%	90%	100%	110%	120%	130%	140%	≥ 150%
Correction factor	1.29	1.16	1.07	1.00	0.94	0.89	0.85	0.81	0.79

Source: Own elaboration based on section 36 h EEG.

[4] This encompasses a very thorough examination of compliance with building and environmental regulation and can take up to 3 months. The approval process is detailed in section 10 BImSchG (2017).

model. Their multi-level optimisation model takes into account long run (investment) decisions on network and generation capacity as well as short run decisions on production, consumption and redispatch. In this framework, there is a trade-off between the concentration of RES capacities at productive sites (which implies high transmission investment and redispatch cost) and RES capacity locations closer to load centres (which comes at lower RES productivity but also lower network expansion and redispatch cost). In one of their scenarios, Grimm et al. (2017) determine the system-optimal RES locations for the current German market design and show that optimal RES locations imply yearly welfare gains of at least €2.6 billion compared to the status quo in Germany.

A precise control of capacity locations via auctions is highly difficult, since capacity expansion and the final locational choice depend on the bidders. However, appropriately designed auction mechanisms can at least steer an allocation in the desired direction and incentivise bidders to build their plants at system compatible sites. We therefore employ the system-optimal onshore wind capacity allocation (MaxW) presented in Grimm et al. (2017) to assess the extent to which it can be achieved with adjustments to the current German wind auction design.

2.4. Alternative auction designs

The previous sections have illustrated several aspects concerning the spatial allocation of RES that are important for auction design. A basic national auction without any locational steering mechanism would allocate RES at the most productive locations and thus yield the cost-minimal remuneration. However, it would ignore additional costs arising from redispatch and prospectively necessary network expansion. Thus, overall system cost from the induced RES expansion would be inefficiently high.

The undesirable concentration of RES at productive sites that are often far from load centres could be mitigated by using mechanisms like the REM, which compensates the disadvantage of less productive sites using correction factors on bids. However, determining the correction factor only based on site productivity and disregarding other important aspects like load proximity limits the effectiveness of the approach. It is only by chance that such a mechanism would induce an almost system-optimal RES allocation.

An alternative approach to ensure a desirable regional distribution is to set regional target capacities. To implement such target capacities, a straightforward approach is to conduct regional instead of national auctions. The corresponding regional target capacities are then tendered in each state or region, thereby exactly implementing the desired regional distribution of capacity. Under appropriate auction rules, within each region, the most productive sites are chosen first. Even though such a regional auction design implements the desired regional distribution of generation capacities, it is vulnerable to market power. In fact, competition within each region is the lower the smaller the regions are, i.e. the more tailored the mechanism is to induce an optimal regional distribution. To ensure a sufficient level of competition in regional auctions, they should be conducted less frequently (e.g. only once a year) and thus, for higher volumes. Note also that separate regional auctions would impose challenges on bidders that consider sites in different regions as substitutes or complements.

As a superior option, a single national auction with predetermined regional capacity quotas or targets could be considered, which allows bidders to place package bids for RES sites in multiple regions. This allows to ensure a sufficient level of competition, both within and between regions. Moreover, such a mechanism exploits the scope for further cost reductions by allowing for package bids that enable bidders to reflect synergies between different projects. Such 'combinatorial' auctions have already been used in industrial procurement, logistics as well as spectrum auctions to increase competition and leverage economies of scale and scope, which could be equally beneficial for the procurement of renewable energies (Bichler et al., 2006; Kokott et al., 2018; Sheffi, 2004; Bichler and Goeree, 2017).

In the remainder of the article we compare the auction formats sketched in this section using numerical experiments on the basis of data for the German electricity market.

3. Model

Let us now introduce the model for our numerical simulations. We first detail our assumptions on project costs and bidder types, before we precisely specify how we implement the analysed auction designs and their respective allocation and pricing rules.

3.1. Project cost

Bidders place bids on individual projects, with a project j describing a wind park of capacity y_j in kW, which can consist of multiple wind power plants at the same location. In order for these projects to operate profitably, bidders need to cover at least their investment costs over the remuneration period of 20 years. Based on regionally differentiated investment costs for onshore wind power plants in $^{€}\!/_{kW}$ and site-specific wind power generation per installed kW, we derive the break-even rate c_j in $^{ct}\!/_{kWh}$ for a project j over the remuneration period of 20 years and divide this by the respective wind efficiency w_j at a project's location in $^{kWh}\!/_{kW}$. Investment costs include all administrative fees and costs for land improvement, while operating costs are assumed to be negligible. To allow for project-dependent variations, we randomly vary the underlying project costs by up to $\pm 5\%$.

The corresponding project capacity y_j is drawn randomly from a PERT distribution with a minimum 0.75 MW, a maximum of 25 MW and a region specific mean. The distributional assumptions are based on the submitted project capacities in 2018 (Federal Network Agency, 2019).[5] For every auction, we consider an individual number of projects per region relative to its size, i.e. at least six projects per region and at most 100 projects in the largest region. The amount of projects is set such that the capacity requirements from both the German network development plan ('Netzentwick-lungsplan', NEP, ÜNB, 2017), which describes the capacity expansion path currently aimed for by German politics and the German auction system, as well as the MaxW allocation (Grimm et al., 2017) can be satisfied. As the largest of the 16 German states, Bavaria would thus be assigned 100 projects. However, due to its strict legal constraints regarding permissible wind park locations, the number of available projects is reduced to 40.[6]

3.2. Bidder types

We assume price taking behaviour of bidders in all our experiments. This allows for a transparent comparison among different auction formats. While renewable energy auctions on a national basis are large markets, we are aware of the fact that strategic behaviour could play a role if separate regional auctions were implemented. Strategic behaviour would be an additional drawback of regional auctions.

In our model there are two types of bidders, namely institutional and BEG bidders. As BEG are small local bidders, they are active in only one region. Each BEG is assigned one project in its respective region, with a capacity $y \leq 18$ MW. Based on the status quo of four auctions per year, we generate up to six BEG bidders per German state and auction, leading to a total of 384 BEG bidders for 2018. Analogously, we consider 120 institutional bidders for 2018, who are each randomly assigned 0–4

[5] We choose a PERT distribution because it fits the empirical data best. PERT is a transformation of the Beta distribution, which allows us to define the shape of the capacity distribution using the minimum, maximum and expected value.

[6] The so-called '10H' rule requires wind power plants to have a distance of at least ten times the wind power plant's hub height to residential areas, see section 82 of the Bavarian Building Law (2007).

projects per auction, i.e. 0–16 projects in 2018. The number of BEG and institutional bidders is based on the 2018 auction data and adjusted such that the capacity requirements in all regions can be met. Bidders participate in every subsequent auction unless all their projects won in earlier auctions.

Since institutional bidders can place bids on multiple projects, they can realise synergies via economies of scale. There are numerous ways how such synergies can come into play, which we try to accommodate by integrating different synergy concepts in our experiments and checking the consistency of our results across these. We account for synergies that arise if a bidder is awarded multiple projects, either subsequently or in a single auction. In particular, we distinguish between synergies for projects within one region (regional synergy), for projects in neighbouring regions (cross-regional synergy), and for all projects in Germany (national synergy). Since BEG are small, local bidders that only place one bid in a single region, they do not realise synergies. An overview of the type as well as number of bidders and synergies is given in Table 3.

Consequently, a project bundle B_i of bidder i can be partitioned into k sets of synergy groups S_k. A synergy group thus contains all projects of a bidder that, depending on the synergy concept, would create scale economies if awarded. Consider e.g. four projects assigned to i from the following regions: R_1, R_1, R_2 and R_3, where R_1 and R_2 are neighbouring regions. In the regional synergy concept, the two projects in R_1 would be in one synergy group, the other two each in a separate one, i.e. $S_1 = \{R_1, R_1\}$, $S_2 = \{R_2\}$, $S_3 = \{R_3\}$. For cross-regional synergies, the groups would be $S_1 = \{R_1, R_1, R_2\}$, $S_2 = \{R_3\}$; while for national synergies all four projects would be in one group. In this respect, the more regions enter a single synergy group, the 'wider' the corresponding synergy concept, i.e. the cross-regional concept is wider than the regional one. The unit cost of each project j in a synergy group S_k is then adjusted by a factor $\lambda \in [0,1]$ to

$$\tilde{c}_j(S_k) = c_j \cdot \left(1 - \left(\lambda \cdot \frac{|S_k| - 1}{|S_k|}\right)\right). \tag{1}$$

λ describes the share of cost savings achieved via synergies. For $\lambda = 0$ there is no synergy. With each additional project in S_k, the marginal effect of synergy decreases. To compute the unit cost of a project bundle, $\overline{c}(B_i)$, we consider the synergy-adjusted unit costs of all individual projects $j \in B_i$ weighted by their respective capacity y_j and wind efficiency w_j, i.e.

$$\overline{c}(B_i) = \frac{\sum_{k=1}^{K} \sum_{j \in S_k} \tilde{c}_j(S_k) \cdot y_j \cdot w_j}{\sum_{j \in B_i} y_j \cdot w_j} \tag{2}$$

An example calculation of synergy-adjusted project and bundle unit costs is given in Appendix B.

3.3. Analysed auction designs

For our analyses, we examine only one year, and we assume that the regulator intends to expand generation capacity proportionately to the final allocation. Based on the respective installed capacity, it is possible to periodically define regional capacity expansion quotas to achieve some target capacity for e.g. 2035. To answer the question whether a certain auction design enables an effective steering of generation capacity expansion according to such quotas, it thus suffices to analyse only one period. More specifically, we do not calibrate our model to most closely match the results of all past auctions.[7] Instead, we aim to provide a sound counterfactual comparison of different auction designs based on their outcome with the administrative parameters given for 2018, the first year during which onshore wind auctions were conducted subject to the current regulatory framework in Germany. In particular, we assess the allocative efficiency of four auction designs, whose allocation mechanisms and pricing rules are described in the following.

Table 3
Overview of bidder types and synergy concepts.

Bidder type	Number of bidders	Projects per bidder	Synergy concept	Synergy levels
BEG	384	1	None	None
Institutional	120	0–16	{regional, cross-regional, national}	[0,0.5]

3.3.1. Single-lot auction designs

In the single-lot auction designs, one bid always corresponds to a single project j.

3.3.1.1. National. The *National* auction design with four auctions a year constitutes our benchmark for comparison, as it implements the cost-minimal outcome.

A bid in the *National* auction design contains the ask price b_j and the capacity y_j for a project j. After all bids are submitted, they are sorted in ascending order by ask price. Bids are accepted as long as the cumulative capacity of accepted bids is smaller than the tendered capacity. Successful BEG bidders receive the remuneration per kWh of the last accepted bid (uniform-price), while institutional bidders receive the remuneration per kWh they asked for (pay-as-bid). This procedure is also shown in pseudo-code below (Algorithm 1).

3.3.2. National REM

The *National REM* auction design essentially describes the current design of the German wind auctions. A total of 2710 MW (the total wind auction volume in 2018 in Germany, see Federal Network Agency, 2019) is auctioned off in four auctions. Bidders place bids according to the REM for a reference site (see section 2.2), which are then tendered as in the *National* auction design. Successful bidders receive a remuneration adjusted by the correction factor corresponding to their bid's relative site quality factor as shown in Table 2.

3.3.3. Regional

The *Regional* auction design exactly implements the regional target capacities of MaxW. In particular, we analyse individual auctions for each of the 16 German states, with one auction per state per year, since the auction volumes would be very small for four auctions per region and year. All 16 regional auctions are conducted simultaneously each year. In each state, the total capacity to be built in 2018 according to MaxW is tendered. Within each region, the most productive and wind-efficient sites are chosen first. The allocation and pricing procedure for *Regional* is identical to the *National* auction design.[8]

3.3.4. Combinatorial auction design

While all previous auction designs allow bids only for single projects, we now propose an auction design that allows to submit package bids: the *Combinatorial* auction design. As in the *Regional* auction design, annual target capacities for every region are set according to MaxW, but auctioned off simultaneously in a single auction. Bids are awarded such that the resulting allocation is as efficient and subsidy-minimising as possible. BEG are local and can only place bids in one region, while institutional bidders can place package bids for projects in multiple regions. In the following, we introduce the corresponding

[7] Since the start of the auction scheme in 2017, the corresponding legal framework has been subject to change in late 2017. Therefore, the available auction data is difficult to interpret. As a matter of fact, only limited information is provided on the auction results. The available data are on an aggregate level or simple averages and thus do not allow an intricate empirical analysis.

[8] This means that for each auction within each German state, institutional bidders are subject to a pay-as-bid auction, while BEG bidders are subject to a uniform auction. Thus in the *Regional* design, all winning BEG from the same state, e.g. Bavaria, receive the same FIP. The resulting FIP for winning BEG can thus differ across states in the *Regional* design.

allocation and pricing rule. When discussing unit prices, we are referring to remuneration per kWh in our context.

For each bidder $i \in I$, we determine whether a bundle of projects $B_i \subseteq P_i$ has been allocated or not by a decision variable $z_i(B_i)$. For an allocated bundle B_i, its unit cost in $^{\text{ct}}\!/_{\text{kWh}}$ is denoted by $\overline{c}(B_i)$, its total capacity in kW at node (i.e. a region) $n \in N$ by $y_n(B_i)$, and its average wind efficiency in $^{\text{kWh}}\!/_{\text{kW}}$ at node n by $w_n(B_i)$. Table 4 summarises the necessary notation.

The sum of the allocated capacities of all accepted bundles must be larger or equal to the tendered capacity d_n for each node $n \in N$ (demand constraint), while each bidder can win at most one of her bundles (supply constraint). The objective of the problem is to select the most efficient projects while minimising excess capacities. To do so, we consider a bundle's unit cost times its capacity. The dual variable of the demand constraint can then be interpreted as the impact of an increase in the demanded capacity d_n of one kW on the objective function, which leads to a price in $^{\text{ct}}\!/_{\text{kWh}}$.

$$\underset{\text{s.t.}}{\text{Min}} \sum_{i \in I} \sum_{B_i \subseteq P_i} z_i(B_i) \cdot \overline{c}(B_i) \cdot \sum_{n \in N} y_n(B_i) \qquad (\mathbf{AP})$$

$$\sum_{i \in I} \sum_{B_i \subseteq P_i} z_i(B_i) \cdot y_n(B_i) \geq d_n \qquad \forall n \in N \qquad (\text{Demand})$$

$$\sum_{B_i \subseteq P_i} z_i(B_i) \leq 1 \qquad \forall i \in I \qquad (\text{Supply})$$

$$z_i(B_i) \in \{0, 1\} \quad \forall i \in I, B_i \subseteq P_i \qquad (\text{Binary})$$

This binary linear program used for the allocation problem of a multi-unit combinatorial procurement auction can be seen as a multidimensional knapsack problem, which is known to be *NP*-hard (Bichler, 2017, p. 100). We can solve the problem sizes in our analysis to near optimality (integrality gap <1%) with standard solvers (Gurobi version 8.1) on commodity hardware within minutes in our experiments.

There is a large literature on pricing in multi-object markets, and more specifically in combinatorial auctions (Bichler, 2017). Let us briefly introduce and discuss some basic considerations for our pricing rule. It is well-known that the Vickrey-Clarke-Groves (VCG) mechanism is incentive-compatible and implements an efficient outcome (Krishna, 2010). However, bidders' payments under VCG are personalised and non-linear. This means that prices for a package of objects can differ from the sum of the prices of individual objects in the package (non-linear), while prices for a package can also vary across bidders (personalised) (Bichler and Goeree, 2017). Moreover, the VCG outcome might not be in the core, i.e. there could be incentives for individuals or coalitions of bidders to deviate (Ausubel and Milgrom, 2006). As argued by Milgrom (2004), discriminatory pricing fails to promote the law of one price and thus may be hard for some people to accept.

Walrasian markets yield linear item-level, anonymous competitive equilibrium prices, but are not incentive-compatible (Hurwicz, 1972). However, Walrasian mechanisms are strategy-proof in the large (Azevedo and Budish, 2019). In other words, such markets are robust to strategic manipulation if there are many bidders who then become price takers. The RES market is relatively large (>100 bidders and projects, Federal Network Agency, 2019, thereby justifying the use of linear and anonymous prices. Such prices are widely used, e.g. in day ahead electricity markets, as well as easy to understand and interpret. Unfortunately, with indivisible goods such as projects in RES auctions, Walrasian prices do not always exist. The types of valuations (e.g. substitutes valuations) for which Walrasian equilibria exist are in fact rather limited (Baldwin and Klemperer, 2019).

In the RES market, linear prices exist if the linear programming relaxation of the allocation problem (AP) is integral. In other words, if the last set of constraints $z_i(B_i) \in \{0, 1\}$ were relaxed to $z_i(B_i) \in [0, 1]$ and the result were still integral, with all $z_i(B_i) \in \{0, 1\}$, then the dual variables of the Demand constraint (p_n) would have a natural interpretation as Walrasian prices. This well-known observation follows from strong duality in linear programming (Bichler, 2017, p. 144ff). The linear programming relaxation of AP is rarely integral, but the integrality gap of the linear programming relaxation, i.e. the difference between the objective function value of the integer program and its relaxation, is small (on average less than 1% in our experiments). This is due to the large number of bidders and bids. We leverage this observation for the pricing rule we introduce in this paper: due to the small integrality gap, also the dual variables of the linear programming relaxation are 'close' to Walrasian prices.

Although the integrality gap is small, market prices cannot simply be derived from the dual linear program, as we show next. To see this, let us first introduce the dual (DAP) of the linear programming relaxation of AP:

$$\underset{\text{s.t.}}{\text{Max}} \sum_{n \in N} d_n p_n - \sum_{i \in I} \pi_i \qquad (\mathbf{DAP})$$

$$\sum_{n \in N} y_n(B_i) p_n - \pi_i \leq \overline{c}(B_i) \sum_{n \in N} y_n(B_i) \qquad \forall B_i \subseteq P_i, \forall i \in I$$

$$\pi_i, p_n \geq 0 \qquad \forall n \in N, \forall i \in I$$

Note that the objective function value of **DAP** is less than or equal to **AP** such that prices will be too low on average, unless we have integrality of the linear programming relaxation. As a result, the prices will be below the bids and bidders can make a loss. To avoid this, we introduce a modified version **MDAP** of the dual linear program to compute prices. MDAP deviates from anonymous prices via a markup only if necessary, and ensures that bidders do not make losses (Individual Rationality, IR). Simultaneously, prices for losing bidders should be below their costs (NOENVY). This approach is akin to pseudo-dual prices as proposed for ascending combinatorial auctions (Bichler et al., 2009).

Let us introduce **MDAP** more formally. If we reverse the inequality in **DAP** and minimise the objective, we get the lowest possible linear and anonymous prices for the winners, such that all winners $W \subseteq I$ can recover their costs, while we also ensure that prices for the losers $L \subseteq I$ are below their costs with $W \cup L = I$ and $W \cap L = \emptyset$. As this is not always possible, we introduce a slack variable, $\delta(B_i)$, for each winner's winning bundle, describing a personalised markup for this bidder. The bidder's payment then exactly covers their costs.

$$\underset{\text{s.t.}}{\text{Min}} \sum_{n \in N} d_n p_n + \sum_{B_i \in W} \delta(B_i) \cdot M \qquad (\mathbf{MDAP})$$

$$\sum_{n \in N} y_n(B_i) w_n(B_i) (p_n + \delta(B_i)) \geq \overline{c}(B_i) \sum_{n \in N} y_n(B_i) w_n(B_i) \qquad (\text{IR})$$

$$\forall i \in W, B_i \subseteq P_i : z(B_i) > 0$$

Table 4
Notation.

Sets	
I	Set of bidders
P_i	Set of projects of bidder $i \in I$
$B_i \subseteq P_i$	Set of projects in a bundle bid of bidder $i \in I$
N	Set of nodes or regions, i.e. states in Germany
Decision variables	
$z_i(B_i)$	Assign bundle B_i to bidder $i \in I$
Parameters	
$y_n(B_i)$	Capacity (size) of bundle B_i in node $n \in N$ (in kW)
$w_n(B_i)$	Average wind efficiency at the generation sites in a bundle B_i (in $^{\text{kWh}}\!/_{\text{kW}}$)
$\overline{c}(B_i)$	Unit cost of bundle B_i in $^{\text{ct}}\!/_{\text{kWh}}$
d_n	Demanded capacity in node $n \in N$ (in kW)
p_n	Unit price in node $n \in N$ (in $^{\text{ct}}\!/_{\text{kWh}}$)

Table 5
Distribution of awarded capacity in 2018 and capacity expansion paths by state and allocation.

State	2018	NEP	MaxW
Schleswig-Holstein (SH)	7.7%	10.2%	0%
Mecklenburg-West Pomerania (MV)	8.8%	16.6%	7.8%
Hamburg (HH)	0%	0%	0%
Bremen (HB)	0.2%	0.1%	0%
Lower Saxony (NI)	12.1%	19.0%	0%
Saxony-Anhalt (ST)	6.2%	8.8%	0%
Brandenburg (BB)	16.9%	5.4%	0%
Berlin (BE)	0%	0%	0.5%
North Rhine-Westphalia (NW)	13.9%	4.9%	11.5%
Saxony (SN)	1.3%	8.1%	4.9%
Thuringia (TH)	3.3%	9.2%	0%
Hesse (HE)	8.0%	2.8%	16.2%
Rhineland-Palatinate (RP)	10.2%	7.2%	6.7%
Saarland (SL)	0.3%	0%	3.5%
Bavaria (BY)	5.2%	0%	35.2%
Baden-Wuerttemberg (BW)	6.7%	7.7%	13.7%
Sum	**100%**	**100%**	**100%**

Note: 0% in the 2018 column means that no capacity was awarded to projects in these states in 2018. 0% in the NEP and MaxW columns indicates that no further capacity expansion is necessary in these states to reach the respective capacity targets, i.e. the optimal capacity in these states subject to the target allocation has already been reached.
Source: Own elaboration based on data from the Federal Network Agency (2019), Grimm et al. (2017) and ÜNB (2017).

$$\sum_{n \in N} y_n(B_i) w_n(B_i) p_n < \overline{c}(B_i) \sum_{n \in N} y_n(B_i) w_n(B_i) \qquad \text{(NOENVY)}$$

$$\forall i \in L, B_i \subseteq P_i$$

$$p_n \geq 0 \quad \forall n \in N$$

$$\delta(B_i) \geq 0 \quad \forall i \in W, B_i \subseteq P_i : z(B_i) > 0$$

The resulting prices from **MDAP** are anonymous and linear for the losers while no winning bidder can make a loss (IR). Markups on the anonymous prices, $\delta(B_i)$, are introduced only when necessary. A penalty term M keeps these deviations as small as possible. In large auctions with small integrality gaps, the price computation in **MDAP** strikes a balance between different design goals.

4. Data and experimental design

In this section we briefly summarise the data used to parametrise our numerical experiments and outline our experimental design.

4.1. Data

We use a variety of historical data sets and data based on authoritative forecasts to calibrate our model. In particular, we employ data on various aspects and areas that are relevant for bidders in wind auctions.

To begin with, we use data on existing and planned renewable energy generation capacity in each German federal state. Information on the installed capacity of onshore wind power plants is taken from Deutsche WindGuard (2018).

In order to generate an amount of projects satisfying the capacity expansion targets in our different treatments, we use data from scenario B 2035 in the NEP (ÜNB, 2017) and on MaxW (Grimm et al., 2019, 2018, 2017, 2016). Both the NEP and MaxW allocations provide capacity targets for 2035, with the planned capacity expansion underlying each allocation corresponding to the difference between the target capacity in 2035 and the current stock. Since we consider only one year in our analysis, we calculate the necessary yearly expansion assuming a linear expansion until 2035.

Table 5 shows the corresponding regional distribution of capacity envisaged according to the NEP (second column) and MaxW (third column). For comparative purposes, the distribution of the capacity awarded in the 2018 auctions is shown in the first column of Table 5.

The auctioned capacity in each bidding round in the *National REM* auction design is defined according to the specifications in section 28 EEG. Based on the resulting annual electricity generation, we determine the respective capacity that needs to be tendered in the *National*, *Regional* and *Combinatorial* auction designs to arrive at the same annual electricity generation.[9] We do so since the reference parameter for RES capacity expansion is generally the annual amount of electricity generation.

Moreover, following Grimm et al. (2017), we account for regional differences in site quality in Germany and within the 16 states by creating 15 classes of technical potential per state to allow for a differentiated simulation of bidding decisions. To generate the classes, we employ data from the Bundesverband WindEnergie (2012). In a next step, we allocate both the existing installed capacity by the end of 2017 and the target capacities for 2018 into the 15 classes in descending order up to their maximum capacity, starting with the best. For our analysis, we assume that within each region new capacity is bid on in the best available class of technological potential first, and in lower classes only once the better ones have reached full capacity. Analogously, we use data on hourly wind power generation per installed kW in $^{kWh}\!/_{kW}$ by state and class of technological potential taken from Grimm et al. (2017).

To adequately model investors' cost structures, we use information on investment costs for wind power plants in 2018 based on Prognos (2013). Since only investment cost data for 2013 and 2035 is given in Prognos (2013), we use linear interpolation to arrive at investment cost values for 2018. Furthermore, we assume spatially differentiated plant configurations and investment costs to account for the varying conditions in the German states. In less windy areas, comparably larger and thus more expensive wind power plants with greater rotor diameter have to be built to generate an amount of electricity per installed kW equal to that in very windy states.

For each of the configurations, we calculate an approximation of the reference yield per installed kW using data on reference yields of comparable existing wind power plants provided by FGW (2017). To be precise, we choose wind power plants whose configurations most closely match those of our four onshore wind categories, extrapolating their reference yields given by FGW (2017) by adjusting for slight differences in hub height or rotor diameter, if necessary. For an overview of the comparative values and wind power plants used for each wind category see Table A.1 in Appendix A. This allows us to simulate the current auction system including the REM. Table 6 provides an aggregated overview.

4.2. Synergies and cost computations

Institutional bidders can develop multiple projects and realise synergies. We are not aware of real-life estimates of such synergies, as bidders tend to be secretive about details of their cost functions. Nevertheless, economies of scale are almost always an issue for large bidders. These can result from e.g. volume discounts, marginal additional administration costs, lower maintenance costs, lower logistic costs, etc. In our study, synergies are a central treatment variable; understanding at which synergy level combinatorial auctions yield lower costs than alternative auction designs that do not consider target capacities is thus of great interest. In our numerical experiments, institutional bidders bid their cost for a project j while they consider synergies arising from previously won projects. They do not speculate on potential synergy effects by winning multiple items in order to avoid making losses.

[9] Though the resulting distribution of wind capacity in the *Regional* and *Combinatorial* auction designs is system-optimal, it is difficult to ascertain how much CO_2 emissions can be avoided. The resulting emissions are largely determined by surrounding circumstances like the extent of the ensuing transmission line expansion and fossil power generation capacities, which are not included in our model. A discussion of the CO_2 emissions resulting from different scenarios and capacity distributions, including MaxW, can be found in Grimm et al. (2017).

Table 6
Investment costs for wind power plants.

	Category	Investment costs [€/kW]	Plant configuration	Reference yield p.a. [MWh/MW]
2018	Onshore Wind 1 (HB, HH, MV, SH)	1, 355	Hub height 95 m, 3 MW, 100 m rotor diameter	2, 321
	Onshore Wind 2 (BB, BE, NI, NW, ST)	1,456	Hub height 105 m, 3 MW, 100 m rotor diameter	2, 376
	Onshore Wind 3 (BW)	1, 630	Hub height 120 m, 2.5 MW, 110 m rotor diameter	3, 915
	Onshore Wind 4 (BY, HE, RP, SL, SN, TH)	1, 732	Hub height 130 m, 2.5 MW, 115 m rotor diameter	4, 065

Source: Own elaboration based on Prognos (2013) and FGW (2017).

Table 7
Preview of experimental design.

Treatment variable	Value
Auction design	{National, National REM, Regional, Combinatorial}
Synergy concept	{regional, cross-regional, national}
Synergy level	{0, 0.1, 0.2, 0.3, 0.4, 0.5}

In the *Combinatorial* auction design, which allows bids on bundles of projects, institutional bidders bid their cost for any possible project bundle $B_i \in P_i$ accounting for possible synergy effects, i.e. they bid $\overline{c}(B_i)$. Bidders in the *National*, *National REM* and *Regional* auction designs consider only synergies from projects that they already won. When placing bids on individual projects, they do not lower their bids by speculating on winning multiple projects. Else, they would risk making losses.

In the *National REM* auction design, as described in Section 2.2, bidders place bids as if their projects were to be built at the reference site. Therefore, each project's individual site-specific wind efficiency w_j in $^{kWh}\!/_{kW}$ is put into proportion with the corresponding reference yield w_j^{REM} from Table 6, i.e. $^{w_j}\!/_{w_j^{REM}}$, which leads to the respective correction factor as shown in Table 2. The final bid results in the break-even costs c_j divided by the correction factor, anticipating that the final remuneration p_j will result from the ask price multiplied by the correction factor.

4.3. Experimental design and focus variables

As described in Section 3.3, we analyse the outcome of four auction designs for three different synergy concepts and various synergy levels. An overview of these treatment variables and their possible combinations is given in Table 7.

To make sure that our results are robust, we analyse ten iterations per treatment combination. With four different auction designs, three synergy concepts, six synergy levels and ten iterations per treatment combination, we thus evaluate a total of 1,800 experimental auctions.[10] More specifically, we assess our results based on the following focus variables:

1. The average remuneration per kWh $\overline{p}$, in ct/kWh
2. The allocative efficiency δ in %, measured as the percentage of capacity allocated to states with capacity demand under MaxW. Naturally, $\delta = 100\%$ in the *Regional* and *Combinatorial* auction designs.
3. The actor diversity η in %, measured as the share of capacity won by BEG bidders.

5. Results

Based on our experiments, we report the effect of the four auction designs on the three primary focus variables: average remuneration per kWh ($\overline{p}$), allocative efficiency (δ) and actor diversity (η). In addition, we report the total payments p.a. (θ, in €m) and the bidders' average cost per kWh ($\overline{c}$) for the interested reader. As stated above, the *National* auction design serves as a benchmark for comparison.

5.1. Synergies

We first compare different auction designs for various synergy levels. A synergy level of 0.2 indicates that a bundle of projects in a synergy group (e.g. a region) can cost up to 20% less than the sum of the individual projects. The more projects are in that group, the closer the cost reduction will be to 20%.

5.1.1. **Result 1** (Synergies)

A given synergy level has a stronger (negative) impact on the remuneration to be paid in the Combinatorial auction than in the National and National REM auction. The lowest impact of synergies on the remuneration is observed in the Regional auction design.

Table 8 presents the effects of our experimental design elements on the average remuneration per kWh $\overline{p}$ obtained from an OLS regression. The wider the synergy concept (regional < cross-regional < national), the lower the remuneration, as synergy groups become larger and the potential for economies of scale increases. However, the difference in effect size between cross-regional and national synergies is small ($-0.3447^{ct}\!/_{kWh}$ vs. $-0.3990\ ^{ct}\!/_{kWh}$), indicating that synergies are most profitably realised for projects in neighbouring or close regions.

The synergy level has the strongest impact on the remuneration per kWh in the *Combinatorial* auction design: an increase of 0.1 in the synergy level decreases the average remuneration by 0.38 $^{ct}\!/_{kWh}$. In comparison, it leads to a decrease of only 0.23 $^{ct}\!/_{kWh}$ in the *National* auction design. Since bidders cannot explicitly account for synergies in the *Regional* auction design, they have no effect on the average price level. This is due to the fact that only one regional auction takes place per year so that bidders cannot account for projects won in earlier auctions when bidding in later auctions. Averaging across synergy concepts, the *Combinatorial* auction design yields the same average price as in the National auction design for a synergy level of 0.59.

5.2. Allocative efficiency and average remuneration per kWh

In a next step, we evaluate allocative efficiency and the corresponding average remuneration in the four auction designs.

5.2.1. Result 2

The National REM auction design yields a higher allocative efficiency compared to the National auction at the expense of a higher average remuneration per kWh. However, the resulting allocation still differs substantially from the desired MaxW. Both the Regional and the Combinatorial auction designs implement MaxW precisely, but at a higher average remuneration. For synergies of 0.4, the Combinatorial auction design yields the average remuneration per kWh of the National REM design. In case of synergies, among all auction formats, the Combinatorial auction design yields by far the lowest bidder margins.

[10] 1,800 auctions =[2(National, National REM) · 4(February, May, August, October) + 2 (Regional and Combinatorial)] · 10(iterations) · 3(synergy structures) · 6(synergy levels)

Table 8
Remuneration effects of design variables.

Dep. variable: avg. remuneration $\overline{p}$	Coef.	SE	t	P>|t|
Intercept	6.5436	0.0280	233.51	0.0000
National REM	0.4597	0.0368	12.49	0.0000
Regional	0.8488	0.0368	23.06	0.0000
Combinatorial	0.8878	0.0368	24.12	0.0000
Cross-regional synergy	−0.3447	0.0180	−19.17	0.0000
National synergy	−0.3990	0.0180	−22.19	0.0000
National × syn. Level	−2.2605	0.0860	−26.29	0.0000
National REM × syn. Level	−2.5355	0.0860	−29.49	0.0000
Regional × syn. Level	−0.0102	0.0860	−0.12	0.9059
Combinatorial × syn. Level	−3.7501	0.0860	−43.62	0.0000
R^2	0.93			
N	720			

Table 9 shows the outcome of the four action designs with respect to the focus variables. We focus on allocative efficiency (δ) and the average remuneration ($\overline{p}$) first. Since there is no information on scale economies for German wind auction bidders in reality, we report our results for both no synergies and cross-regional synergies with $\lambda = 0.2$ and $\lambda = 0.4$.

Allocative efficiency, average remuneration and bidders' costs increase when applying the REM. Without synergies, 89% of the subsidised capacity in *National REM* are allocated to regions with a positive capacity demand under MaxW. This is 40 percentage points more than in the *National* auction design. As synergy increases, these shares as well as their difference decrease, the latter from 40 to 28 ($\lambda = 0.2$) and 25 ($\lambda = 0.4$) percentage points. This is mostly due to a decline in allocative efficiency in the *National REM* design, since a higher cross-regional synergy supports a wider distribution of projects. For a graphical illustration of the resulting allocations, see Figs. C.1 and C.2 in Appendix C.

Meanwhile, the average remuneration per kWh in the *National REM* design is 7.2% ($\lambda = 0.0, 0.2$) and 5.4% ($\lambda = 0.4$) higher than in the *National* auction design. Note that while the REM slightly increases allocative efficiency, bidders lack incentives to search for efficient sites when it is applied.

Irrespective of the synergy level, allocative efficiency is at 100% in both the *Regional* and *Combinatorial* auction designs, as we only allow bidding in regions with positive capacity demand under MaxW. But this comes at a cost: without synergies, the average remuneration per kWh in both designs increases by 14.4% and 14.7% compared to the *National* design, respectively, which can be considered a surcharge for forcing the system-optimal allocation of MaxW. Note that, considering no synergies, the *Regional* auction design leads to slightly lower prices than the *Combinatorial* auction designs. The reason for this is that the *Regional* auction design only minimises prices while the *Combinatorial* auction design minimises prices and overcapacity, leading to lower total cost (θ).

Table 9
Comparison of auction design outcomes.

Auction design	Synergy level (λ)	$\overline{p}(\frac{ct}{kWh})$	$\delta(\%)$	$\eta(\%)$	$\theta(m$ p.a.)	$\overline{c}(\frac{ct}{kWh})$
National	0	6.25	49	24	366	6.11
National REM	0	6.70	89	27	389	6.60
Regional	0	7.14	100	19	421	7.11
Combinatorial	0	7.17	100	18	417	7.13
National	0.2	5.81	46	15	332	5.43
National REM	0.2	6.23	74	11	355	5.86
Regional	0.2	7.14	100	19	413	6.52
Combinatorial	0.2	6.34	100	5	362	6.14
National	0.4	5.20	45	8	294	4.51
National REM	0.4	5.48	70	7	308	4.79
Regional	0.4	7.14	100	19	407	5.95
Combinatorial	0.4	5.50	100	5	310	5.04

However, the average remuneration per kWh in the *Combinatorial* auction design becomes similar to that of the currently applied *National REM* design in the presence of synergies, and remains only 9.1% ($\lambda = 0.2$) and 5.8% ($\lambda = 0.4$) higher than in the *National* auction design. In other words, for only moderate synergies, the *Combinatorial* auction design implements the system-optimal allocation without any surcharge compared to the current German auction design, while maintaining incentives to search and bid on the most efficient sites.

It is also notable that in case of synergies the *Combinatorial* auction design reduces bidders' margins substantially as compared to all other auction designs: from 0.38 (0.37, 0.62) ct/kWh in the *National* (*National REM, Regional*) auction design to 0.20 ct/kWh for $\lambda = 0.2$ and from 0.69 (0.69, 1.19) ct/kWh to 0.46 ct/kWh for $\lambda = 0.4$.

5.3. Bidder diversity

5.3.1. Result 3

The higher the synergies, the more institutional bidders benefit from economies of scale and the less capacity is allocated to BEG bidders.

The fifth column of Table 9 reports the share of awarded capacity won by BEG bidders, $\eta(\%)$. Without synergies, this ranges from as much as 27% in the *National REM* auction design to 18% in the *Combinatorial* design. With increasing synergy levels, institutional bidders gain a competitive advantage; consequently, the share of successful BEGs decreases substantially and generally lies below 10% for $\lambda = 0.4$. Since the *Regional* auction design is unaffected by synergies, the share of BEG bidders stays at 19%.

5.4. Computational cost

5.4.1. Result 4

All auction formats can be computed in seconds with realistic problem sizes. It takes significantly longer to compute the allocation and remuneration per kWh for the Combinatorial auction design than for the other auction designs, on average 5 seconds.

The average time required to solve the allocation and corresponding pricing problems in each auction design is provided in Table 10. Note that we are solving real-world problem sizes. The computation times would not constitute a practical problem.

5.5. Summary

Forcing the system-optimal allocation of MaxW comes at the cost of an average increase in remuneration of about 14% compared to the *National* auction design. This changes in the presence of synergies: for $\lambda = 0.2$ and $\lambda = 0.4$ the average remuneration is only 9.1% and 5.8% higher, respectively, all the while maintaining the system-optimal capacity allocation of MaxW.

The percentage of capacity won by BEG bidders decreases with stronger synergy effects, as institutional bidders gain a comparative advantage. Note that maintaining a steady and moderate share of successful BEG bidders can also be a policy goal. This could easily be implemented with additional allocative constraints in the *Combinatorial* auction design.

6. Conclusion

Many countries are using auctions to determine the remuneration for RES, which is however often accompanied by a high concentration of renewable energy power plants at very productive sites far-off the main load centres. To counteract these tendencies, we introduce a combinatorial auction design that allows to implement regional target capacities, provides a simple pricing rule and maintains a high level of competition between bidders by permitting package bids.

The aim of this paper was to assess and evaluate the impact of four different RES auction designs on allocative efficiency and subsidy

Table 10
Computation times.

Auction design	Mechanism	Mean (std) in sec.
National	Allocation	0.013 (0.000)
National REM	Allocation	0.014 (0.000)
Regional	Allocation	0.021 (0.000)
Combinatorial	Allocation	4.47 (0.045)
National	Pricing	0.003 (0.000)
National REM	Pricing	0.005 (0.000)
Regional	Pricing	0.019 (0.000)
Combinatorial	Pricing	0.816 (0.054)

payments by means of extensive numerical experiments. Based on the case of onshore wind auctions in Germany, we compare the current nationwide auction design with the REM to a simple nationwide auction design, a regional and our proposed combinatorial auction design.

We find that for only moderate synergies, the *Combinatorial* auction design implements the system-optimal wind capacity allocation presented by Grimm et al. (2017) without considerably increasing the average remuneration per kWh compared to the current German auction design, while maintaining incentives to search and bid on the most efficient sites. Grimm et al. (2017) estimate the potential savings resulting from a system-optimal allocation of RES in Germany to be at least €2.6 billion a year for a 2035 scenario. Current cost experiences for redispatch and feed-in management measures often range above a billion Euro per year and are mainly caused by the high concentration of onshore wind power plants in northern Germany, far from the main load centres in southern Germany. This indicates a high potential for savings resulting from our proposed Combinatorial auction design.

The prices resulting from the Combinatorial auction are linear and anonymous for each region whenever possible, while minimal personalised markups on the linear prices are applied only when necessary to prevent winning bidders from making losses. At the same time, prices are set such that no losing bidder would want to produce at those prices. Due to the size of the problem instances (i.e. the tendered capacity and number of bids), the personalised markups are minimal.

Combinatorial auctions come at the cost of computational complexity for the auctioneer since the allocation problem that needs to be solved is an NP-hard combinatorial optimisation problem. In our experiments, we show that realistic problem sizes can be solved in seconds due to the large number of relatively small bidders. For bidders, combinatorial auctions are strategically simpler than having to bid in a sequence of auctions. In particular, institutional bidders can bring their scale economies to bear with package bids, which reduces costs.

Overall, the combinatorial auction design proposed in this paper is a viable alternative to location-specific auction mechanisms like the German REM. Furthermore, it constitutes a candidate design for renewable energy auctions not only in Germany, but also in other countries worldwide where auctions are used to support the expansion of renewable energy capacity.

Acknowledgements

The authors acknowledge support by the Bavarian State Government through the Energy Campus Nuremberg (EnCN), the International Doctoral Program “Evidence-Based Economics” of the Elite Network of Bavaria the Kopernikus project "SynErgie" (BMBF, Federal Ministry of Education and Research, Germany) by Martin Bichler CRC TRR 154 subproject B08 by Veronika Grimm). We would like to thank Christian Soelch as well as seminar participants in Munich and Nuremberg for many fruitful discussions on the topic of this paper.

Appendix A

Table A.1
Comparative values for reference yield calculation.

Category	Parameter	Configuration	Comparative value	Plant type	Reference yield p.a. [$\frac{MWh}{MW}$]
1	Rotor diameter (m)	100	100.5	eno 100	3,195.36
	Hub height (m)	95	99		
	Nominal capacity (MW)	3	2.2		
2	Rotor diameter (m)	100	100.5	eno 100	3,195.36
	Hub height (m)	105	99		
	Nominal capacity (MW)	3	2.2		
3	Rotor diameter (m)	110	112	Vestas V112	2,965.89
	Hub height (m)	120	119		
	Nominal capacity (MW)	2.5	3.3		
4	Rotor diameter (m)	115	112	Vestas V112	3,192.87
	Hub height (m)	130	140		
	Nominal capacity (MW)	2.5	3.3		

Source: Categories based on Prognos (2013), comparative values taken from FGW (2017).

Appendix B

Example 1. A bidder i has $P_i = \{P1, P2, P3, P4, P5, P6\}$ projects, represented by tuples of (region, cost). Those are: ('BW',10), ('BW',10), ('BY',10), ('RP',10), ('BB',10), ('BE',10). Assuming there are no synergies, the unit cost for each project is: $\widetilde{c}_j(S_k) = c_j = 10\,\frac{ct}{kWh}, \forall j \in P_i$ and all partitions S_k of P_i.

Assume synergies are considered to be $\lambda = 0.5$ for projects within a region. The sets of projects in the same synergy groups are $S_1 = \{P1, P2\}, S_2 = \{P3\}, S_3 = \{P4\}, S_4 = \{P5\}, S_5 = \{P6\}$. When winning S_k, the respective unit cost for each project in S_1 is given by: $\widetilde{c}_j(S_1) = 10 \cdot (1 - 0.5 \cdot \frac{1}{2}) = 7.5\,\frac{ct}{kWh}$. The unit cost for all other $k > 1$ is $\widetilde{c}_j(S_k) = 10\,\frac{ct}{kWh}$.

Appendix C

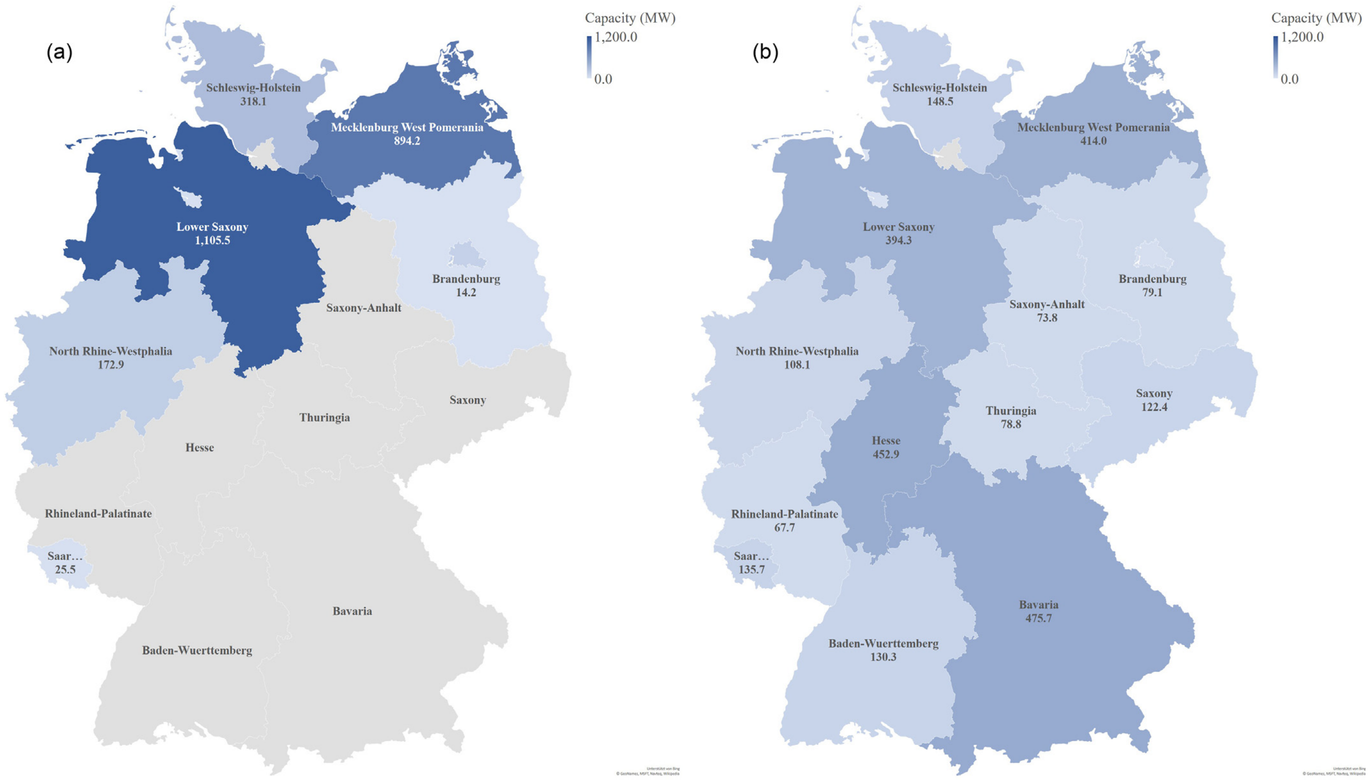


Fig. C.1. Capacity allocation in the (a) *National* and (b) *National REM* auction design for cross-regional synergies of $\lambda = 0.2$.

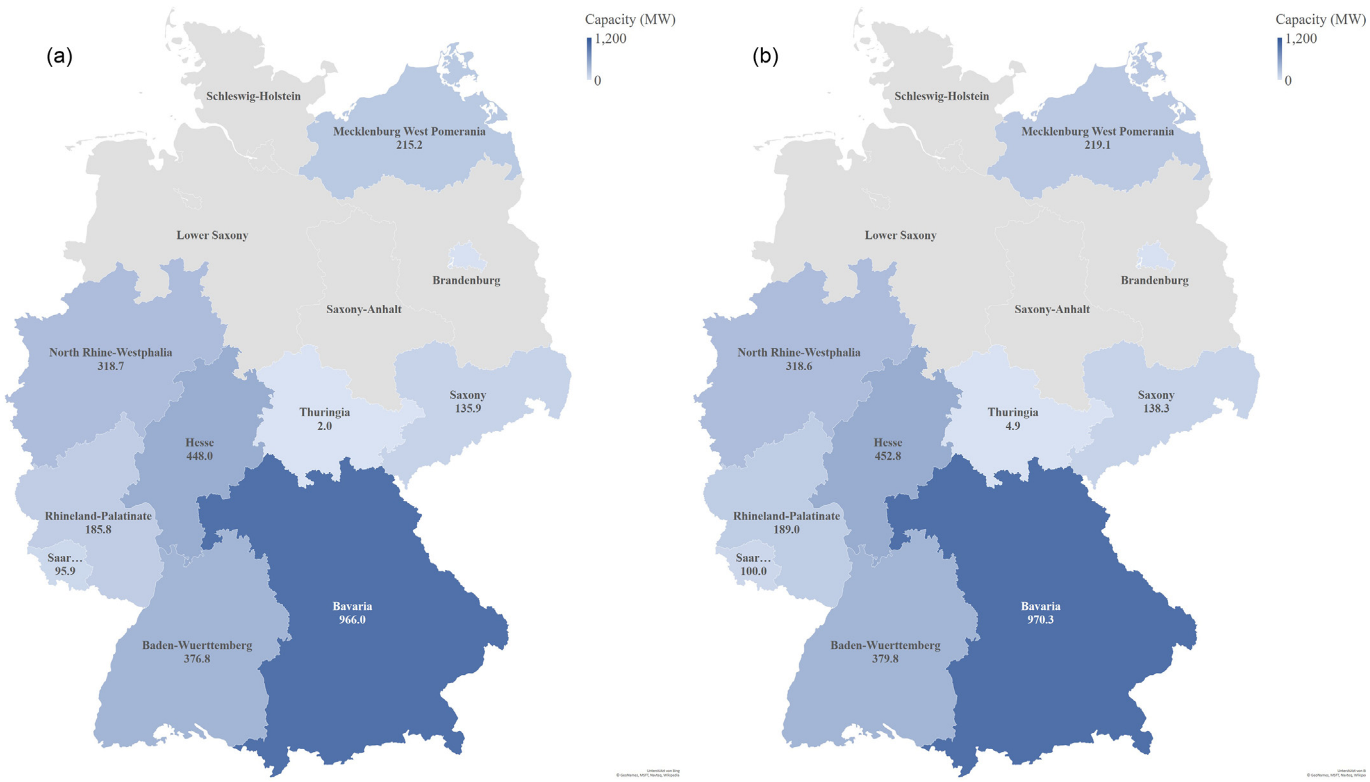


Fig. C.2. Capacity allocation in the (a) *Regional* and (b) *Combinatorial* auction design for cross-regional synergies of $\lambda = 0.2$.

References

Ackermann, T., Andersson, G., Söder, L., 2001. Distributed generation: a definition. Electr. Power Syst. Res. 57, 195–204.

Adomavicius, G., Curley, S.P., Gupta, A., Sanyal, P., 2012. Effect of information feedback on bidder behavior in continuous combinatorial auctions. Manag. Sci. 58, 811–830.

Amado, M., Poggi, F., Ribeiro Amado, A., Breu, S., 2017. A cellular approach to net-zero energy cities. Energies 10, 1826.

Anatolitis, V., Welisch, M., 2017. Putting renewable energy auctions into action – an agent-based model of onshore wind power auctions in Germany. Energy Policy 110, 394–402.

Anaya, K.L., Pollitt, M.G., 2015. The role of distribution network operators in promoting cost-effective distributed generation: lessons from the United States of America for Europe. Renew. Sust. Energ. Rev. 51, 484–496.

Ausubel, L.M., Milgrom, P., 2006. The lovely but lonely Vickrey auction. In: Cramton, P., Shoham, Y., Steinberg, R. (Eds.), Combinatorial auctions Chapter 1. (Pp. 17–40) Boston.

Azevedo, E.M., Budish, E., 2019. Strategy-proofness in the large. Rev. Econ. Stud. 86, 81–116.

Baldwin, E., Klemperer, P., 2019. Understanding preferences: "demand types", and the existence of equilibrium with indivisibilities. Econometrica 87, 867–932.

Barroso, L.A., Street, A., Granville, S., Pereira, M.V., 2011. Offering strategies and simulation of multi-item iterative auctions of energy contracts. IEEE Trans. Power Syst. 26, 1917–1928.

Bavarian Building Law, 2007. Bayerische Bauordnung (BayBO) in der Fassung der Bekanntmachung vom 14. August 2007 (GVBl. S. 588, BayRS 2132–1-B), zuletzt durch § 1 des Gesetzes vom 10. Juli 2018 (GVBl. S.523) geändert.

Benz, T., Dickert, J., Erbert, M., Erdmann, N., Johae, C., Katzenbach, B., Glaunsinger, W., Müller, H., Schegner, P., Schwarz, J., 2015. Der zellulare Ansatz: Grundlage einer erfolgreichen, regionenübergreifenden Energiewende (Technical Report VDE).

Bichler, M., 2017. Market Design: A Linear Programming Approach to Auctions and Matching. Cambridge University Press, Cambridge.

Bichler, M., Goeree, J.K., 2017. Handbook of Spectrum Auction Design. Cambridge University Press, Cambridge.

Bichler, M., Davenport, A., Hohner, G., Kalagnanam, J., 2006. Industrial procurement auctions. In: Cramton, P., Shoham, Y., Steinberg, R. (Eds.), *Combinatorial Auctions* Chapter 23. MIT Press, Boston, pp. 593–612.

Bichler, M., Shabalin, P., Pikovsky, A., 2009. A computational analysis of linear Price iterative combinatorial auction formats. Inf. Syst. Res. 20, 33–59.

Bichler, M., Gupta, A., Ketter, W., 2010. Research commentary - designing smart markets. Inf. Syst. Res. 21, 688–699.

Bichler, M., Hao, Z., Adomavicius, G., 2017. Coalition-based pricing in ascending combinatorial auctions. Inf. Syst. Res. 28, 159–179.

BImSchG, 2017. Gesetz zum Schutz vor schädlichen Umwelteinwirkungen durch Luftverunreinigungen, Geräusche, Erschütterungen und ähnliche Vorgänge. Bundes-Immissionsschutzgesetz in der Fassung der Bekanntmachung vom 17. Mai 2013 (BGBl. I S. 1274), das zuletzt durch Artike.

Contreras, J., Candiles, O., De La Fuente, J.I., Gomez, T., 2001. Auction design in day-ahead electricity markets. IEEE Trans. Power Syst. 16, 88–96.

Cozzi, P., 2012. Assessing Reverse Auctions as a Policy Tool for Renewable Energy Deployment. Tufts University Medford, Technical Report The Fletcher School.

Cramton, P., 2010. How best to auction natural resources. In: Daniel, P., Keen, M., McPherson, C. (Eds.), The Taxation of Petroleum and Minerals. Routledge, London and New York, pp. 289–316.

EEG, 2017. Gesetz für den Ausbau erneuerbarer Energien (Erneuerbare-Energien-Gesetz). Erneuerbare-Energien-Gesetz vom 21. Juli 2014 (BGBl. I S. 1066), das zuletzt durch Artikel 5 des Gesetzes vom 13. Mai 2019 (BGBl. I S. 706) geändert worden ist.

European Commission, 2014. Guidelines on State Aid for Environmental Protection and Energy 2014–2020.

Federal Ministry for Economic Affairs and Energy, 2016. What is actually the reference yield model?

Federal Ministry for Economic Affairs and Energy, 2018. Windenergie an Land.

Federal Network Agency, 2019. Statistiken zum Ausschreibungsverfahren für Windenergieanlagen an Land. Technical Report Last accessed June 12, 2019.

Federal Network Agency, 2019a. Ausschreibungen zur Ermittlung der finanziellen Förderung von Windenergieanlagen an Land.

Federal Network Agency and Federal Cartel Office, 2019. *Monitoringbericht 2018.* Technical Report Monitoring report in accordance with section 63(3) i. c. w. section 35 EnWG and section 48(3) i. c. w. section 53(3) GWB. Status as of February 8, 2019, Bonn.

Fernandez-Blanco, R., Arroyo, J.M., Alguacil, N., 2014. Network-constrained day-ahead auction for consumer payment minimization. IEEE Trans. Power Syst. 29, 526–536.

Fitch-Roy, O., Woodman, B., 2016. Auctions for Renewable Energy Support in the United Kingdom: Instruments and lessons learnt. Technical report AURES report D4.1-UK.

Fördergesellschaft Windenergie und andere Dezentrale Energien (FGW), 2017. Referenzertragsverfahren.

Förster, S., 2016. Small-scale PV Auctions in France: Instruments and lessons learnt. Technical report AURES report D4.1 - FR.

Gallachóir, B.P.Ó., Bazilian, M., McKeogh, E.J., 2009. Wind-energy policy development in Ireland: A critical analysis. In: Stachan, P., Lal, D., Toke, D. (Eds.), Wind Power and Power Politics. Routledge, New York, pp. 128–155.

Gerlach, A.-K., Pape, C., Stratmann, P., 2015. Regional market value factors for wind power. Wind Energy Report Germany 2014. Fraunhofer IWES, Bremerhaven and Kassel, pp. 85–89.

Grashof, K., 2013. Herausforderungen bei Ausschreibungsverfahren für Onshore-Windenergie und PV. In: Leprich, U., Grashof, K., Guss, H., Klann, U., Weber, A., Zipp, A. (Eds.), *Stromsystem-Design: das EEG 2.0 und Eckpfeiler eines zukünftigen Regenerativwirtschaftsgesetzes.* Saarbrücken: IZES.

Grimm, V., Martin, A., Schmidt, M., Weibelzahl, M., Zöttl, G., 2016. Transmission and generation investment in electricity markets: the effects of market splitting and network fee regimes. Eur. J. Oper. Res. 254, 493–509.

Grimm, V., Sölch, C., Zöttl, G., 2017. Regionalkomponenten bei der EE-Vergütung (Technical Report Report on behalf of the German Monopoly Commission).

Grimm, V., Rückel, B., Sölch, C., Zöttl, G., 2020. The Impact of Market Design on Transmission and Generation Investment in Electricity Markets, Forthcoming Energy Economics.

Grimm, V., Rückel, B., Sölch, C., Zöttl, G., 2019. Regionally differentiated network fees to affect incentives for generation investment. Energy 177, 487–502.

Guo, Z., Koehler, G.J., Whinston, A.B., 2012. A computational analysis of bundle trading markets design for distributed resource allocation. Inf. Syst. Res. 23, 823–843.

Güsewell, D., 2016. Die Lösung könnte in einer Quote liegen.

Hauser, E., Kochems, J., 2014. Ausschreibungsmodelle für Wind Onshore: Erfahrungen im Ausland (Technical Report IZES Saarbrücken).

Hauser, E., Weber, A., Zipp, A., Leprich, U., 2014. Bewertung von Ausschreibungsverfahren als Finanzierungsmodell für Anlagen erneuerbarer Energienutzung (Technical Report IZES Saarbrücken).

Heer, K.-D., Langniß, O., Wiser, R., Lewis, J., Hamrin, J., 2007. Promoting Renewable Energy Sources in Portugal: Possible Implications for China. ZSW Stuttgart, June.

Held, A., Ragwitz, M., Gephart, M., De Visser, E., Klessmann, C., 2014. Design Features of Support Schemes for Renewable Electricity. Technical Report Ecofys Utrecht, Netherlands.

Hurwicz, L., 1972. On informationally decentralized systems. In: McGuire, C., Radner, R. (Eds.), Decision and Organization. North-Holland, Amsterdam.

Ibrahim, H., Ghandour, M., Dimitrova, M., Ilinca, A., Perron, J., 2011. Integration of wind energy into electricity systems: technical challenges and actual solutions. Energy Procedia 6, 815–824.

IRENA and CEM, 2015. Renewable Energy Auctions - a Guide to Design (Technical Report).

IRENA and GWEC, 2012. 30 Yeary of Policies for Wind Energy (Technical Report Abu Dhabi).

de Jager, D., Rathmann, M., 2008. Policy Instrument Design to Reduce Financing Costs in Renewable Enery Technology Projects (Technical Report by Ecofys).

Jürgens, C.A., 2017. *Harmonisierung von Windenergie- und Netzausbau vor dem Hintergrund der Engpassbewirtschaftung. Eine Untersuchung der politischen Steuerungsmechanismen im Spannungsfeld zwischen Versorgungssicherheit und kostengünstiger Umsetzung der Energiewende in Deutsc.* Ph.D. thesis. Helmut-Schmidt-Universität Hamburg.

Kardakos, E.G., Simoglou, C.K., Bakirtzis, A.G., 2013. Short-term electricity market simulation for pool-based multi-period auctions. IEEE Trans. Power Syst. 28, 2526–2535.

Kitzing, L., Wendring, P., 2015. Auctions for Renewable Support in Denmark: Instruments and lessons learnt. Technical report AURES report D4.1-DK.

Klemperer, P., 2004. Auctions: Theory and Practice. Princeton University Press, Princeton.

Kokott, S., Levchenko, S.V., Rinke, P., Scheffler, M., 2018. First-principles supercell calculations of small polarons with proper account for long-range polarization effects. New J. Phys. 20, 33023.

Kreiss, J., Ehrhart, K.-M., Haufe, M.-C., 2017. Appropriate design of auctions for renewable energy support - Prequalifications and penalties. Energy Policy 101, 512–520.

Krishna, V., 2010. Auction Theory. 2nd ed. Elsevier, Burlington, San Diego, London.

Maurer, L.T., Barroso, L.A., 2011. Electricity Auctions: An Overview of Efficient Practices. The World Bank, Washington DC.

McLean, A., McGovern, P., Donelly, K., 2007. Ireland. International Energy Law and Taxation Review. 10 pp. 207–210.

Milgrom, P.R., 2004. Putting Auction Theory to Work. Cambridge University Press, Cambridge.

Mora, D., Kitzing, L., Rosenlund Soysal, E., Steinhilber, S., del Río, P., Wigand, F., Klessmann, C., Tiedemann, S., Amazo, A., Welisch, M., Kreiß, J., Fitch-Roy, O., Woodman, B., 2017. Auctions for Renewable Energy Support - Taming the Beast of Competitive Bidding (Technical Report AURES Final report).

Negri, A., 2015. Auctions Design for the Renewable Energy Aid: The Italian Experience. In *Presentation at IRENA Workshop*, Renewable Energy Auctions Design and Best Practice.

Peña Cabra, I.A., 2014. *Retrospective and prospective analysis of policy incentives for wind power in Portugal.* Ph.D. thesis. Carnegie Mellon University.

Prognos, 2013. Entwicklung von Stromproduktionskosten: Die Rolle von Freiflächen-Solarkraftwerken in der Energiewende. Technical Report accessed March 5, 2018.

del Río, P., 2017. Designing auctions for renewable electricity support. Best practices from around the world. Energy for Sustainable Development 41, 1–13.

del Río, P., Linares, P., 2014. Back to the future? Rethinking auctions for renewable electricity support. Renew. Sust. Energ. Rev. 35, 42–56.

del Río, P., Haufe, M.-C., Wigand, F., Steinhilber, S., 2015. Overview of design elements for RES-E auctions. Technical Report AURES Report D2, 2.

Rosen, C., Madlener, R., 2013. An auction design for local reserve energy markets. Decis. Support. Syst. 56, 168–179.

Sheffi, Y., 2004. Combinatorial auctions in the procurement of transportation services. Interfaces 34, 245–252.

Tang, L., Wu, J., Yu, L., Bao, Q., 2017. Carbon allowance auction design of China's emissions trading scheme: a multi-agent-based approach. Energy Policy 102, 30–40.
Tiedemann, S., Förster, S., Wigand, F., 2016. Auctions for Renewable Energy Support in Italy: Instruments and lessons learnt. Technical report AURES report D4.1 - IT.
Toke, D., 2015. Renewable energy auctions and tenders: how good are they? International Journal of Sustainable Energy Planning and Management 08, 43–56.
ÜNB (2017). Netzentwicklungsplan Strom 2030, Version 2017, Zweiter Entwurf der Übertragungsnetzbetreiber. May 2, 2017.
de Vos, R., Klessmann, C., 2014. How to Design a Successful Auction for Renewable Energy Projects.
Wigand, F., Förster, S., Amazo, A., Tiedemann, S., 2016. Auctions for renewable energy support: lessons learnt from international experiences. Technical Report AURES Report D4, 2.
WindEnergie, Bundesverband, 2012. Potenzial der Windenergienutzung an Land (Technical Report).
WindGuard, Deutsche, 2017. Status des Windenergieausbaus an Land in Deutschland (Technical Report Report on behalf of the Bundesverband Windenergie and VDME Power Systems).
WindGuard, Deutsche, 2018. Status des Windenergieausbaus an Land in Deutschland (Technical Report Report on behalf of the Bundesverband Windenergie and VDMA Power Systems).
Zhang, L., Li, Z., Wu, C., 2014. Randomized auction design for electricity markets between grids and microgrids. SIGMETRICS Performance Evaluation Review 42, 99–110.

5 Computing Approximate Bayes-Nash Equilibria through Neural Self-Play

Peer-Reviewed Workshop Paper

Title: Computing Approximate Bayes-Nash Equilibria through Neural Self-Play

Authors: S. Heidekrüger, P. Sutterer, M. Bichler

In: Workshop on Information Technologies and Systems 2019

Abstract: Understanding market dynamics means understanding and predicting the behaviour of the market participants. Nash equilibria have proven to be an effective means in this regard. Unfortunately, computing equilibria in a complete information or Bayesian game is computationally hard. We introduce a learning rule based on neural networks that we call Neural Self-Play. This rule is able to compute approximate Nash equilibria for many normal form games as well as for incomplete-information games with continuous type- and action-space, i.e., sealed bid single-item auctions. Leveraging GPU hardware architecture, which allows for parallelized computation of large matrices, Neural Self-Play finds approximate Bayes-Nash equilibria in first-price sealed bid auctions with 10 players within 10s of minutes.

Computing Approximate Bayes-Nash Equilibria through Neural Self-Play

Stefan Heidekrüger* Paul Sutterer† Martin Bichler‡

Department of Informatics
Technische Universität München

Abstract

Understanding market dynamics means understanding and predicting the behaviour of the market participants. Nash equilibria have proven to be an effective means in this regard. Unfortunately, computing equilibria in a complete information or Bayesian game is computationally hard. We introduce a learning rule based on neural networks that we call Neural Self-Play. This rule is able to compute approximate Nash equilibria for many normal form games as well as for incomplete-information games with continuous type- and action-space, i.e., sealed bid single-item auctions. Leveraging GPU hardware architecture, which allows for parallelized computation of large matrices, Neural Self-Play finds approximate Bayesian Nash equilibria in first-price sealed bid auctions with 10 players within 10s of minutes.

1 Introduction

Market design has received increasing attention in the information systems literature (Bichler et al. 2010). For market designers, it is important to understand equilibrium behavior of market participants to predict market outcomes and potential strategic problems. While early literature on general equilibrium theory focused on competitive equilibria and assumed players to be non-strategic price takers, auction theory assumes strategic agents and uses the Nash equilibrium concept to study the price formation process (Nash et al. 1950). More precisely, auction theory models auctions as Bayesian games and analyzes the Bayes-Nash equilibria of players.

Unfortunately, for many markets we do not know the Bayes-Nash equilibrium strategy. For example, Bayes-Nash equilibrium strategies for simple combinatorial first-price sealed-bid auction are still unknown, except for restricted environments (Kokott et al. 2019). Different assumptions on the common prior distribution, the risk aversion of the players, or the number

*stefan.heidekrueger@in.tum.de
†paul.sutterer@tum.de
‡bichler@in.tum.de

of players and objects all play a role, and the analytical derivation of equilibrium strategies can be very challenging, often without a closed-form solution if at all possible.

In this paper, we introduce Neural Self-Play (NSP), a method that numerically derives Bayes-Nash equilibria. In experiments, we focus on environments where we know the analytical solution and show that NSP closely approximates the analytical equilibrium strategy. This bears the promise that it can provide such a solution for markets where we cannot derive analytical solutions. While earlier literature either stems from artificial intelligence or game theory, equilibrium computation becomes increasingly important as a tool in market design and other areas of information systems research. This also contributes to the overall theme of the workshop: markets for policy making and sustainability.

1.1 Related Literature

Nash equilibria (NE) are a central solution concept in non-cooperative game-theory. Informally, in a Nash equilibrium no agent has an incentive to deviate, given the current behaviour of all other agents. Therefore, once a NE is found, it is a stable state. However, finding NE is hard. Actually, it is known to be PPAD complete already for 2-player normal-form games (Daskalakis et al. 2009) and it is hard to approximate (Rubinstein 2016).

There exist a number of learning rules which try to find NE, two of the most frequently used are Fictitious Play (FP) (Brown 1951) and Smooth Fictitious Play (SFP), a variant of the first. The idea of FP is an iterative pre-play process in which each player plays a best response to the opponents' expected play, based on past observations. FP applies to games of complete information, such as normal form games, as well as to games of incomplete information. While FP works fine for many games, its direct application fails whenever a game has continuous type- and action-space, as in auctions.

The problem of numerically computing approximate NE in auctions with continuous type- and action-spaces has previously been studied by Bosshard et al. (2017). Bosshard et al.'s algorithm is shown to compute verifyable approximate equilibria in the general setting. They discretize and transform the Bayesian auction game into a normal form game and compute a pointwise best response. Afterwards, they apply a (smoothed) best response update in the original continuous game by interpolating the discrete solution such that it guarantees an upper bound on the utility loss. While their method is shown to converge, the complexity of calculations in the required discretization grows exponentially with the number of players and thus becomes intractable for games with many players or multidimensional type or action spaces.

1.2 Contributions

In this study, we introduce Neural Self-Play (NSP), a learning rule implementing players' strategies as neural networks, and using evolutionary strategies to update the networks parameters. We first test the algorithm's performance on normal form games of complete information with discrete type- and action-spaces. The algorithm performs similar to FP and SFP and is able to find pure Nash equilibria (PNE) as well as mixed Nash equilibria (MNE) in empirical frequencies. While FP and SFP are only able to work with discrete type- and action-spaces, NSP also works with continuous settings. We test the algorithm on games of incomplete information with continuous type- and action-spaces, i.e. on sealed bid single-item auctions. NSP is able to find approximate Bayes-Nash equilibria (BNE) in all performed experiments within minutes. It is able to find BNE for settings with many bidders and scales well even for an increasing number of parameters.

The remainder of this paper is structured as follows: First, in Section 2, we introduce preliminaries as well as NSP and related learning rules. We then present empirical results of applying NSP to normal form and auction games in Section 3 before concluding with a summary of our findings in Section 4.

2 Methodology

In this study, we apply different learning rules for finding Nash equilibria (NE), namely Fictitious Play (FP), Smooth Fictitious Play (SFP), Mixed Fictitious Play (MFP) which are well-studied tabular methods. In addition, we introduce a new algorithm for equilibrium learning based on neural networks that we call Neural Self-Play (NSP). Before we describe these learning rules, let us briefly introduce a few terms.

Games in normal form (complete information, discrete type- and action space) are defined by a tuple: $G = (N, \mathcal{A}, u)$ where $N = \{1, ..., n\}$ describes the set of players; $\mathcal{A} = \mathcal{A}_1 \times ... \times \mathcal{A}_n$ describes the set of action profiles, with $\mathcal{A}_i$ being the set of actions available to player i; and $u = (u_1, ..., u_n)$ is the joint utility function where $u_i : \mathcal{A}_i \to \mathbb{R}$ describes the payoff (utility) function for each player.

Games of incomplete information are described by a quintuple: $G = (N, \mathcal{A}, V, p, u)$. N and $\mathcal{A}$ are as above, with $\mathcal{A}_i$ potentially being continuous sets $\mathcal{A}_i \subset \mathbb{R}$; $V = V_1 \times ... \times V_n$ is the set of type profiles. At the beginning of the game, each player i is informed of her own type $v_i \in V_i$ only (private information). Just as $\mathcal{A}_i$, the V_i are (potentially continuous) subsets of $\mathbb{R}$.[1] $p(v)$ defines a prior probability distribution over type profiles that is assumed to be common knowledge. The payoff (utility) function is now determined by $u_i : \mathcal{A} \times V_i \to \mathbb{R}$, i.e. players' utilities depend on all players' actions but only their own type.

In each game, after receiving the private type information, each player i chooses her strategy according to some (possibly stochastic) strategy $\pi : V_i \to \Delta\mathcal{A}_i$ that maps to a probability distribution over possible actions.[2] All the learning rules described here have in common that the underlying game is played repeatedly—in theory indefinitely—while players observe each other's behavior and adjust their strategies π_i ("learning") in order to ultimately find an equilibrium in the game, i.e. a state where no player can improve their own expected utility by changing their strategy π_i any further. Throughout this paper, we denote by the index $-i$ a profile of types, actions or strategies for all players but player i.

2.1 Fictitious Play

FP was first introduced by (Brown 1951). It can be seen as a process of pre-play by each player to learn more about the game's dynamics. In FP, each player starts with initial beliefs about the other players' strategies and updates these beliefs based on the observations of

[1] Private information may also be multidimensional, but we restrict ourselves to the scalar setting here.

[2] When π_i is known to be deterministic, i.e. return an action a_i with probability 1, we will use the following abuse notation: $\pi_i(v_i) = a_i$.

played actions throughout the process. At each step every player i computes her expected utility u_i for any possible action in $\mathcal{A}_i$, given the current beliefs of opponents play σ_{-i}, and chooses the action $a_i \in \mathcal{A}_i$ that maximizes it, i.e. plays a best response:

$$a_i = \arg\max_{a \in \mathcal{A}_i} \mathbb{E}\left[u_i(a, \sigma_{-i})\right]$$

After each round, players update their beliefs about other players' strategies using Bayesian updating. As the actual play can only converge to pure Nash equilibria (PNE) due to the way actions are chosen, it oscillates in games with only mixed Nash equilibria (MNE). However, the empirical distribution of historical actions may still converge in such games (Fudenberg and Levine 1999, p.42 - 45) and is thus usually considered when speaking about convergence of FP. While FP does not converge in general (Shapley 1964), it has been shown to converge for some general settings such as constant sum games (Robinson 1951) or games that are solvable through iterated elimination of strictly dominated strategies (Nachbar 1990). For details on convergence guarantees of FP, we refer the interested reader to any text book on game theory, e.g. Fudenberg and Levine (1999).

2.2 Smooth Fictitious Play

Smooth Fictitious Play (SFP) is based on FP but differs in that SFP does not deterministically play a best response, but adds randomness to the decision process. In our implementation this is achieved by applying the softmax function to the expected utilities of each action and sampling an action according to the resulting probability distribution. We further apply a temperature parameter τ that controls the level of smoothing, i.e. the degree of indifference between actions. For $\tau \to \infty$, players will be completely indifferent between actions; as $\tau \to 0$, the players probability of playing the best response action approaches 1. Usually, τ is initialized with 1 and decreases with each step. The probability of player i to play an action a, given the beliefs of opponents playing σ_{-i}, is then given by:

$$\mathbf{Pr}\left(a_i \,|\, \sigma_{-i}\right) = \frac{e^{\frac{u_i(a_i, \sigma_{-i})}{\tau}}}{\sum_{r_i \in A_i} e^{\frac{u_i(r_i, \sigma_{-i})}{\tau}}},$$

where we dropped the expectation around $u_i(\cdot, \sigma_{-i})$ for ease of notation.

SFP can be motivated in multiple ways, among them are: the randomization represents private information about the utility function of a player; and the introduction of randomization allows agents to be less exploitable. In contrast to FP, the actual play in SFP (or the probability for the actions according to players' strategies π) is in principle able to converge to MNE (Fudenberg and Levine 1999, p.131 - 156).

2.3 Mixed Fictitious Play

Mixed Fictitious Play (MFP) is an adjustment of SFP in which players do not sample an action but can "play" mixed strategies that are observed by others. This adjustment makes MFP purely fictitious, i.e. a mind experiment, since players cannot actually play a probability but would have to decide on an action in practice. The advantage is faster convergence due to lack of noise introduced by sampling. This method is thus only suited for finding potential NE of a game but not a method for players to learn to reach the equilibrium strategy through repeated playing of an actual game.

2.4 Neural Self-Play

We propose Neural Self-Play (NSP) as an alternative iterative learning rule that is applicable both to normal form games and continuous-action continuous-type Bayesian games. In NSP, we model players' strategies using neural networks. In each step, players consider their opponents to be stationary in their current strategy (as opposed to updating beliefs over historical play as in FP and variants). The general idea of NSP is that players apply a small update to their neural network parameters θ that will lead to an improvement in utility.

The canonical way of implementing this idea would be applying a gradient ascent algorithm via backpropagation. In fact, this method is called Policy Gradients in (single-agent) reinforcement learning and has been previously studied in multi-agent normal form games where it is called Infinitesimal Gradient Ascent (IGA, Singh et al. 2000; Bowling and Veloso 2002). However, in the following, we demonstrate that this approach fails in the setting of auctions and instead propose to use an alternative training algorithm based on Evolutionary Strategies, before introducing the specific model architectures that we use in this study.

2.4.1 Infinitesimal Gradient Ascent

In Infinitesimal Gradient Ascent (IGA), each player adjusts their own strategy in the direction of the gradient of their utility function when considering opponents fixed at their current strategies[3]:

$$\pi_i^{t+1} := \pi_i^t + \alpha \nabla_{\pi_i} u_i(\pi_i^t, \pi_{-i}^t)$$

In 2x2 normal form games, this simple learning rule has been shown to either converge to a NE or end up in cycling behaviour where each player's average utilities converge to those in a NE (Singh et al. 2000). However, IGA relies on knowledge of the analytical joint gradient dynamics and assumes that the joint utility function is differentiable everywhere. This makes the learning rule unsuitable for continuous-type, continuous-action Bayesian games as these can involve nontrivial discontinuities as we discuss below. To rectify this, our approach differs from IGA mainly in the way gradients are computed, particularly in two aspects:

On the one hand, IGA assumes analytical knowledge of—or an efficient way to compute with arbitrary precision—the global gradient vector field in each step. This assumption, however, becomes impracticable in infinite information-state spaces, as no closed-form description might exist or be known. We thus forgo this assumption and instead rely on stochastic estimation of the gradients: In each iteration, we play a *batch* of games, i.e. draw a batch of valuation profiles from the players' prior distributions and calculate players' current strategy utilities in each of the valuation profiles. Due to the parallel nature of these calculations, we can leverage modern hardware accelerators such as GPUs to perform these batched operations at no additional cost in computation time. We then aim to calculate the gradients for this stochastic joint utility function with respect to each player, which in expectation will approximate the gradient dynamics of the full Bayesian Game.

[3]The sceptical reader might wonder how ∇_{π_i} is defined. π_i could either be a tabular vector of action-probabilities, a parametrized function, etc.; the gradient should be understood with respect to the respective representation determining π_i. We use abuse of notation here to illustrate the concept in a general way.

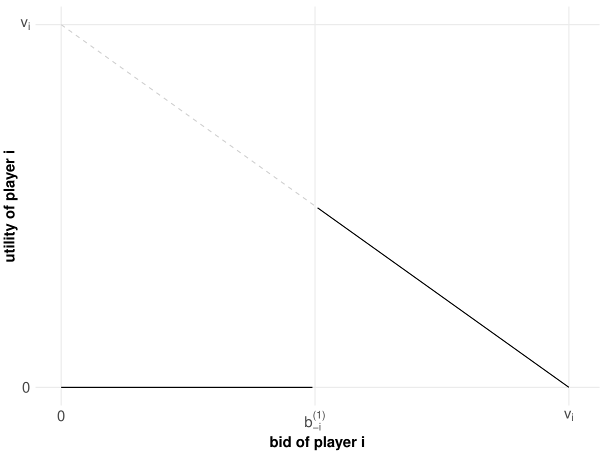


Figure 1: Utility function $u_i(b_i)$ in First Price Sealed Bid Auction for stationary opponent bids b_{-i} with highest opponent bid $b_{-i}^{(1)}$. For a given current bid b_i, the gradient $\nabla u_i(b_i)$ will be zero whenever the player is not winning the item, and negative whenever she is. Thus, when all players update their strategies using gradients, they will eventually all bid zero, as the winner in each round learns to bid less while the losers do not change their strategy.

2.4.2 Evolutionary Strategy Pseudo-gradients

On the other hand, even when available, the exact gradients on this sample may not lead to proper learning, so we rely on pseudo-gradients computed via an Evolutionary Strategy algorithm instead. Exact gradients are problematic because for a fixed valuation and opponent strategy profile (v_i, π_{-i}), player i's utility may be discontinuous in her action. Clearly, such a discontinuity is relevant to playing optimally, but neither the left-sided nor the right-sided derivatives will contain information about its presence, as outlined in Figure 1. The standard method of training neural networks, stochastic gradient ascent (SGA)[4] via backpropagation, calculates exact gradients with respect to the training data[5]; thus, using backpopagation in the multi-agent setting is simply an implementation of IGA on neural network strategies, leading to the problems described above and making it unsuitable in our setting.

Recently, Evolutionary Strategies (ES) have been proposed as an alternative to backpropagation for gradient estimation in neural networks and applied with some success in reinforcement learning (Salimans et al. 2017). In ES, the parameter vector θ of the model is perturbed randomly P times, for example by adding P i.i.d. zero-mean, σ^2-variance Gaussian noise terms ε_p. The resulting P perturbed neural networks are then evaluated with respect to their "fitness" F_p and the model is ultimately updated using a weighted average of the P noise vectors ε_p with more desirable perturbations being weighted higher than less desirable ones: $\theta^{t+1} = \theta^t + \alpha \frac{1}{P} \sum_{m=1}^{P} F_m \varepsilon_p$. While Salimans et al. (2017) mainly motivate this alternative update with the need for large scale parallelization across CPU clusters and computational deficiencies of backpropagation, the method also exhibits an important property that is crucial in our context:

[4]In the Machine Learning and Nonlinear Programming literature the method is commonly known as stochastic gradient descent (SGD). Nevertheless, we will use the maximization formulation here.

[5]Unless impaired by numerical precision.

The ES pseudo-gradient is in expectation identical to the analytical infinitesimal gradient for $\sigma \to 0$; however, in pactice, a small but strictly positive value for σ is used. The resulting *finite* perturbations solve the problem of inconsistent gradient signals at discontinuities of the utility: If an agent is 'barely' losing an auction, a small perturbation resulting in a higher bid will also result in the agent winning the auction, thus providing a positive pseudo-gradient signal. We therefore propose using neural networks trained via ES rather than backpropagation in the multi-agent continuous-action setting whenever the marginal utility functions may not be differentiable or even continuous in action-space.

In our implementation, we extend the basic ES algorithm from Salimans et al. (2017) with two common practices from Reinforcement Learning and Optimization by (a) using the player's utility in the previous iteration as a baseline parameter to reduce variance in the fitness function and (b) replacing the pseudo-gradient update with a momentum update in order to smoothen the learning trajectories. A complete description of Neural Self-Play with Evolutionary Strategy training is given in Algorithm 1.

2.4.3 Representing Strategies: The policy network

In NSP, each agent's strategy is given by a *policy model* that maps her types v_i to (a distribution over) actions and that is represented by a neural network with a parameter vector θ: $\pi_i(\cdot) = \pi_{i,\theta}(\cdot)$.

In the normal form game setting, we implement the policy model as follows: Since we have complete information, there are no information sets and no structure for the neural network to learn. Thus, the policy model for each player consists of a single weight vector representing logits for each possible action, $\theta \in \mathbb{R}^{|\mathcal{A}_i|}$. The logits are then normalized by a softmax function to achieve a vector that can be interpreted as probability distribution: $\mathbf{Pr}\,(a) = \frac{e^{\theta_a}}{\sum_{j\in\mathcal{A}_i} e^{\theta_j}}$. This can be interpreted as a no-hidden-layer feed-forward neural network with a constant scalar input of 1 and an output layer with weight vector θ, no bias parameters and a softmax activation function. A schematic of this network is shown in Figure 2a. Actions are then sampled from the resulting distribution.

In the Bayesian, continuous-type-and-action setting, we instead restrict ourselves to deterministic policies: The input to the neural network will be a vector representing the player's private information, the output will be a vector in action space. In the setting of sealed bid single-item auctions, both the input (private valuations v_i) and outputs (bids b_i) happen to be scalars. The deterministic action is then given by $b_i = \pi_{i,\theta}(v_i)$. To map inputs to outputs, we may use an arbitrary neural network architecture; in this study, we restrict ourselves to fully-connected feed-forward networks with two hidden layers, which were sufficient to yield desired results. Advanced network architectures such as recurrent neural networks or attention mechanisms may be required to extend this technique to settings with temporal structure (such as ascending auctions), but we leave this investigation to future work.
For the reported experiments we use *SeLU* activations (Klambauer et al. 2017) in the hidden layers, and a ReLU activation function in the ouput layer. While the ReLU activation in the output layer fulfills a structural role in ensuring non-negative bids, SeLU was chosen in the hidden layers because we found it to be most robust in producing good results. The network architecture used in auction games is illustrated in Figure 2b.

Algorithm 1: Neural Self-Play with Evolutionary Strategy training

Input: players $i \in [N]$ with initial policy $\pi_i^0 := \pi_{i,\theta_i^0}$, defined by a model architecture and initial parameter vector θ_i^0;
batch size K; learning rate schedule $(\alpha^t)_{t\geq 1}$; friction parameter $\beta \in [0, 1)$;
ES population size P; ES noise stddev σ

For each player, initialize momentum buffer $m_i^0 = 0$
for $t := 1, 2, \dots$ **do**
  For each player i, sample a batch of valuations $v_{k,i}$ for $k \in [K]$
  Calculate joint utility in current strategy profile:

$$u^{t-1} := \frac{1}{K} \sum_k u\left(\pi^{t-1}(v_k)\right)$$

  for each player i **do**
    Sample P perturbations of player i's current policy model:

$$\tilde{\pi}_p := \pi_{i,\tilde{\theta}_p}, \qquad \text{with } \tilde{\theta}_p := \theta_i^{t-1} + \varepsilon_p, \quad \varepsilon_p \sim \mathcal{N}(0, \sigma^2 I) \text{ iid. } \forall p \in [P]$$

    Evaluate the fitness of perturbations by playing a batch vs current opponents:

$$F_p := \frac{1}{K} \sum_k u_i\left(\tilde{\pi}_p(v_{k,i}), \pi_{-i}^{t-1}(v_{k,-i})\right) - \underbrace{u_i^{t-1}}_{\text{baseline}}$$

    Calculate ES pseudo-gradient as fitness-weighted perturbation noise:

$$\nabla^{ES} := \frac{1}{\sigma^2 P} \sum_p F_p \varepsilon_p$$

    Perform a momentum update on the current policy:

$$m_i^t := \beta m_i^{t-1} + \nabla^{ES}$$
$$\theta_i^t := \theta_i^t + \alpha^t m_i^t$$
$$\pi_i^t := \pi_{i,\theta_i^t}$$

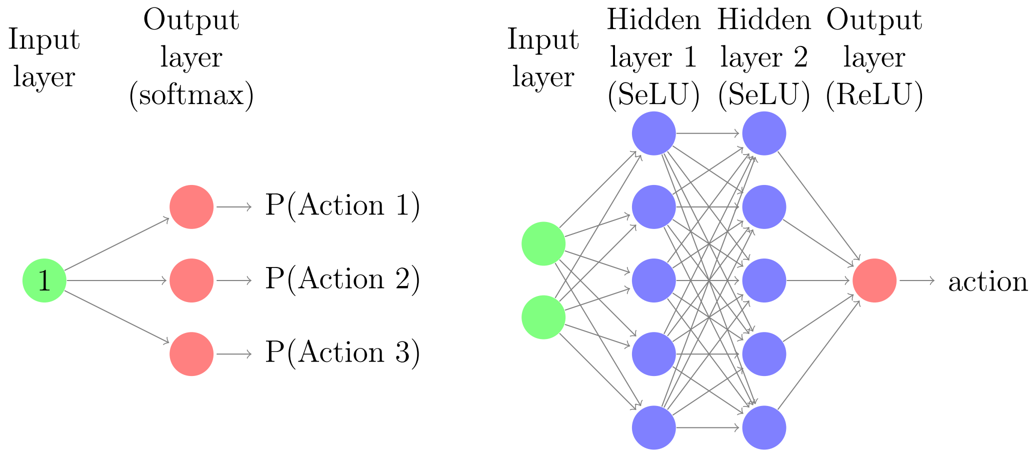
  end
end

3 Empirical Results

We study the learning rules above in two settings, namely complete information normal form games and incomplete information single item sealed-bid auctions.

3.1 Normal Form Games

We consider a number of very common normal form games, starting with 2 players and 2 actions, namely Prisoners Dilemma (PD), Battle Of the Sexes (BoS), and Matching Pennies (MP). We also consider a game with 3 players and 2 actions, namely the Jordan-Game (JG) that has been considered a challenge for FP and its variants (Jordan et al. 1993). We run 10 replications of each game with randomly drawn initial beliefs that are identical for each learning rule. In each replication, each learning rule performs 5000 (learning) steps. The



(a) Normal Form Game Policy Network

(b) Auction Game Policy Network

Figure 2: neural network architectures used for normal form games and auctions.

temperature τ (SFP and MFP) is initialized with 1, updated every 10 steps with 0.9 times the previous value and held constant at a minimum of 0.2. NSP is performed with a batch size of $K = 2^{10}$, ES noise parameter $\sigma = 5$, and $P = 10$ ES perturbations per step.

3.1.1 2 Player, 2 Action

The following games are described by $N = \{1, 2\}$ and $\mathcal{A}_1 = \mathcal{A}_2 = \{1, 2\}$. Due to restricted space, we do not present each payoff matrix but only name the Nash equilibria (NE). In PD, the only NE is both players playing action 2 (PNE). In BoS, there are two PNE (both players play 1 or both players play 2) and one MNE, where player 1 has a 60% probability of playing action 1 and a 40% probability of playing action 2, while the probabilities are reversed for player 2. In MP, the unique MNE is that both players have a 50% probability of playing action 1. Figure 3 illustrates the learning process for these three games and the four learning rules. Since there are only two actions and the behaviour of player 2 is very similar to that of player 1, we display only the actual probability of player 1 playing action 1 at any learning step (columns 1-3). The fourth column displays the empirical distribution of historical probabilities of player 1 playing action 1. Let us describe the results now.

FP (row 1) quickly converges to PNE in PD and BoS (column 1-2) while it oscillates between actions in scenarios of only MNE, here MP (column 3), as described in the previous section. However, the empirical distribution (column 4) converges to the MNE. On the other hand, SFP (row 2) has a tendency of playing mixed and therefore takes about 150 steps to finally play PNE in PD and BoS (column 1-2). However, in games of only MNE, SFP converges to the MNE in actual play and not only in the empirical distribution (column 3 and 4). Actually, Fudenberg and Kreps (1993) established global convergence to a Nash distribution in 2×2 games with a unique mixed-strategy equilibrium. MFP (row 3) generally behaves like SFP, however converges much faster and much smoother, especially in MP (column 3 and 4). NSP (row 4) behaves similarly to both FP and SFP. In PD and BoS (column 1 and 2), it is similar to SFP and needs 200 steps in both to converge to the PNE. In MP (column 3), it is similar to FP and actual play cycles between the actions. However,

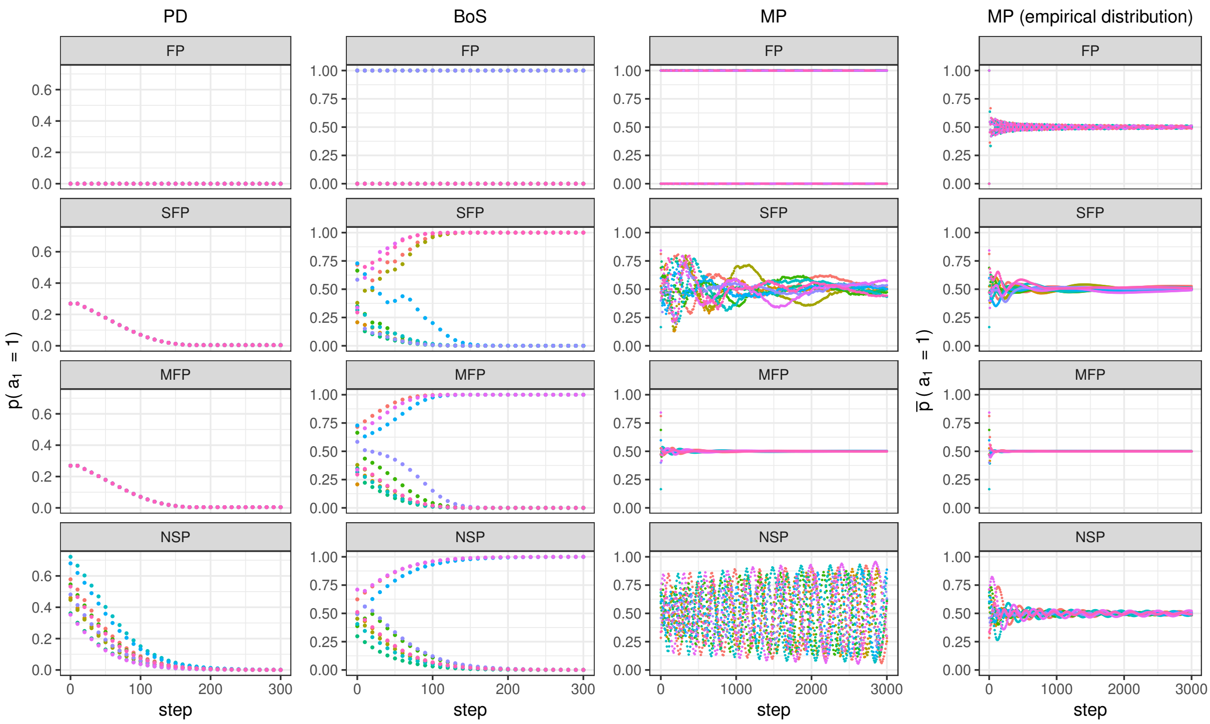


Figure 3: Learning process of player 1 to play action 1 for the four learning rules on three common 2 player 2 actions normal form games. The actual probability to play action 1 at each learning step is shown in column 1-3 and the empirical distribution of historical probabilities in column 4.

this cycling is much smoother than in FP. As in FP, the empirical distribution (column 4) converges to the MNE.

3.1.2 3 Player, 2 Action

While all previous games considered only 2 players, the Jordan-Game (JG) is defined by $N = \{1, 2, 3\}$ and $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}_3 = \{1, 2\}$. In this game, player 1 wants to choose an action different to that of player 2 ($u_1 = 1$, else $u_1 = 0$), player 2 wants to choose an action different to that of player 3 ($u_2 = 1$, else $u_2 = 0$), and player 3 one that is different to player 1 ($u_3 = 1$, else $u_3 = 0$). The only NE is for all players to play each action with a probability of 0.5 (MNE). Figure 4 shows the probability of actual play (row 1) and the empirical distribution of historical probabilities (row 2) for each learning rule (columns 1-4) in the JG.

FP (column 1) oscillates in actual play (row 1). However, in contrast to MP, even the empirical distribution (row 2) does not converge but cycles around the equilibrium. Only in one repetition the empirical distribution is perfectly in the MNE. Here, the initial beliefs are such that each player beliefs all other players play action 1, and therefore each player plays action 2. In the next step, each player updates their beliefs and now beliefs all other players play action 2, and therefore each player plays action 1, etc.. Note that while all players play the MNE in the empirical distribution, the actual payoff is 0 at all times. SFP (column 2)

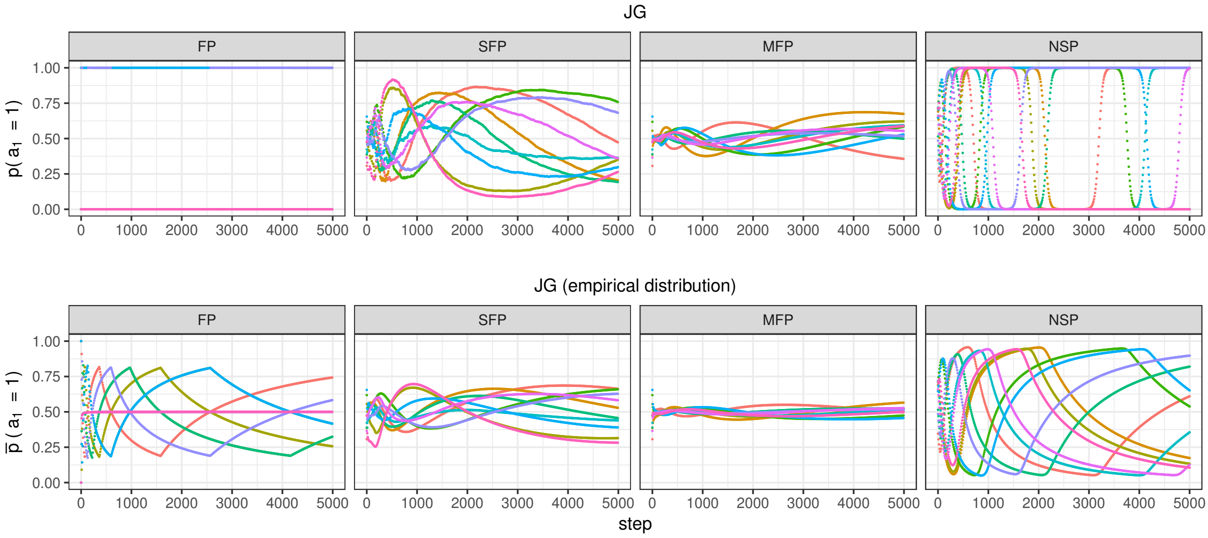


Figure 4: Convergence of actual and historical probability of player 1 to play action 1 for the four learning rules on Jordan Game

does not converge either but both actual play (row 1) and the empirical distribution (row 2) cycle around the equilibrium. This has also been shown by Benaım and Hirsch (1999). The results are the same for MFP (column 3). However, the cycles around the equilibrium are much closer here which is in line with previous results of SFP and MFP. NSP (column 4) performs similar to FP but with even larger cycles of the empirical distribution around the equilibrium. Both would not be suited to find equilibria in this game.

3.2 Single-Item Sealed-Bid Auctions

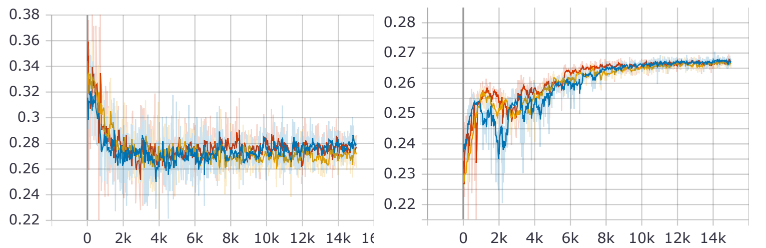
We study NSP behaviour in two types of auctions: First Price Sealed Bid (FPSB) auctions and Second Price Sealed Bid Auctions (also called Vickrey auctions).[6] The latter is well known to be incentive compatible, thus bidding truthfully constitutes a BNE for any combination of valuation distributions. NSP learned a close approxmation to the truthful strategy after just a few 100s of iterations in all Vickrey settings we performed (uniform and normal distributed types with up to 10 players). We thus omit detailed quantitative results for Vickrey auctions for brevity and instead focus on the more challenging case of FPSB auctions.

In FPSBs, analytical Bayes-Nash equilibria (BNE) are known for n players with arbitrary but symmetric prior valuations (Menezes and Monteiro 2005) as well as for 2 players with asymmetric uniform valuation distributions with a stronger and a weaker player (Plum 1992). We ran experiments in the symmetric settings with uniform and normal distributed valuations for 2, 3, 5 and 10 players each. In this setting, we take advantage of the symmetry and implement NSP with model sharing, i.e. symmetric agents share a common parameter vector θ. In this way, the for-loop in line 5 of algorithm 1 will only have to be computed once

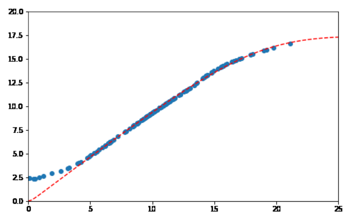
[6]We also implemented NSP with policy gradient training via backpropagation and found that, in fact, the problematic behaviour described in Section 2.4.1 always emerges in practice.

Auction	**Valuation Priors**	**n**	**runs**	**iters**	**runtime (mins)**	**Utility in BNE**	**Utility NSP self play**	**Utility NSP vs BNE**	**Relative utility loss vs BNE (%)**
First Price	**Uniform Symmetric**	2	5	2000	7.05 (0.13)	1.667	1.659 (0.007)	1.665 (0.001)	0.083 (0.038)
		3	5	2000	7.83 (0.14)	0.833	0.827 (0.008)	0.832 (0.001)	0.208 (0.100)
		5	5	3000	13.9 (0.10)	0.333	0.335 (0.006)	0.332 (2. e-4)	0.280 (0.079)
		10	3	3000	15.5 (0.15)	0.091	0.100 (0.006)	0.089 (0.001)	2.289 (1.114)
	Uniform Asymmetric	2	6*	5000	18.1 (0.32)	weak: 0.969	0.901 (0.025)	0.958 (0.005)	1.160 (0.518)
						strong: 5.069	5.102 (0.046)	5.033 (0.007)	0.699 (0.149)
	Normal Symmetric	2	5	5000	10.7 (1.13)	2.779	2.639 (0.074)	2.758 (0.013)	0.778 (0.560)
		3	5	5000	24.5 (1.39)	1.401	1.390 (0.057)	1.398 (0.018)	0.876 (1.313)
		5	6*	10000	67.6 (5.89)	0.668	0.676 (0.013)	0.667 (0.001)	0.103 (0.149)
		10	3	15000	56.5 (1.56)	0.269	0.275 (0.013)	0.267 (0.002)	0.861 (0.565)

Table 1: Results of Neural Self-Play in FPSB auctions. For each metric, we report the mean (and standard-deviation) of multiple runs as indicated. See Table **??** for experiment hyper-parameters. **Note:* In the uniform asymmetric and normal symmetric 5-player settings, one run each failed to learn (at least one player bidding constant zero). In these cases, reported results are calculated over the remaining 5 runs.



(a) Trajectory of NSP utility in self-play (left) and vs the analytical Bayes-Nash equilibrium (middle) over three runs of 15k iterations each. Opaque lines have been smoothed exponentially, actual values indicated with transparency.



(b) Sampled bids (y-axis) vs valuations (x) of learned strategy (yellow run in (a), 99.2% efficiency). BNE indicated in red.

Figure 5: NSP in 10-player FPSB with symmetric-normal valuation priors

in each time step, giving considerable speedups, especially in settings with many players. In the asymmetric setting, the weak and strong player naturally have distinct models. Each experiment was run on a single Nvidia Geforce RTX 2080Ti GPU with batch sizes chosen as large as possible such that the experiment would fit into GPU-memory.

In all of these settings, we measure players' utilities in self-play as well as when unilaterally playing against the known analytical BNE in each observation. We observe convergence of the players' utilities to those achieved in the BNE for both notions of utility in all considered settings: With rudimentary manual hyperparameter tuning, we achieve more than 97.5% efficiency in all 9 FPSB settings and more than 99% in all but two. Detailed FPSB results are presented in Table 1. Figure 5 shows selected learning behaviour and resulting policy in the settings of 10 player symmetric normal valuations.

In asymmetric settings, an interesting phenomenon can sometimes be observed in early training: Initially, one player i will often randomly play a bid strategy that dominates all other players (i.e. i wins all auctions in the batch) but that is nevertheless below the equilibrium bid level. i will then adjust to bid less globally, while other players increase

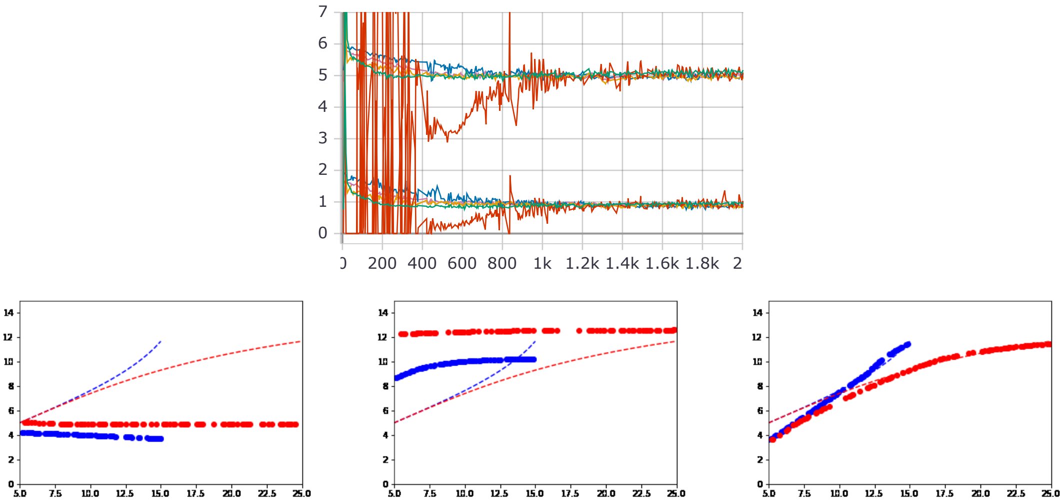


Figure 6: (top) NSP utilities (self-play) in first 2000 iterations of 5 repetitions in asymmetric valuation setting showing both the strong (BNE utility: 5.07) and weak player (0.97). (bottom) NSP strategies corresponding to the red trajectory on (top). Bidding strategies of weak (blue) and strong(red) player as learned after 200, 400 and 2000 iterations (f.l.t.r.).

their bid when they get close to winning. This is precisely the behaviour where in IGA, the 'losing' agent will fail to adjust their bids upwards, resulting in all players bidding 0 after a while. In NSP with ES, however, we can see that due to the upward correction of the losing bidder, the level where winning and losing players 'flip' adjusts upward over time until it reaches the equilibrium level. An example of this can be seen in Figure 6: The erratic behaviour in the red trajectories corresponds to this phenomenon and results in oscillations between achieved utilities much higher than in equilibrium (when 'winning') and 0 (when 'losing'). Ultimately, the level of bids where these flips happen rise to amounts similar to the equilibrium at which point players learn to coordinate, each player wins a fraction of the auctions in each batch.

As expected in deep learning settings, we find that NSP behaviour is sensitive to the choice of hyperparameters in terms of runtime and performance. In our experiments, hyperparameters were chosen and tuned manually and should by no means be considered optimal for their respective settings. In particular, the choice of learning rate α and friction β were found to be of high importance. A too high learning rate (and β) can lead to oscillations around the optimum without convergence. In extreme cases large update steps even lead to a player submitting all-zero bids in one iteration. This results in a behaviour similar to a 'dead ReLU' in backpropagation, where ES can no longer produce valid pseudo-gradient information and the player will bid constant-zero in all following iterations. On the other hand, small learning rates naturally lead to very slow convergence, especially in the setting with many (5, 10) players. A detailed overview of hyperparameters used in the experiments can be provided upon request.

It should further be noted, that even for small ϵ, an ϵ-BNE might be arbitrarily distant from an exact BNE in type-action space (compare Bosshard et al. 2017). As such, we plotted players' policy functions over time for inspection. We often see behaviour where agents do not conform to the equilibrium strategy for low valuations, particularly in settings with high number of players (compare Figures 5b and 6 (bottom right)). This results from the fact that in equilibrium, a player with a low random valuation will almost never win an auction, and even if she does, the utility gained will be minuscule. In fact, analysis of learning behaviour shows that agents learn the correct bid-level for their highest valuation levels very quickly, then fine-tune the shape of the policy.

4 Conclusion

Nash equilibria are popular means to predict market participants' behaviour and predict market outcomes. Unfortunately, computing Nash equilibria is extremely difficult, in fact PPAD complete. In this study, we propose a new learning rule based on neural networks that we call Neural Self-Play. First, we show that this learning rule can compete with common learning rules like Fictitious and Smooth Fictitious Play in normal form games. While these common learning rules require discrete type- and action-space, we show that Neural Self-Play is able to find Bayes-Nash equilibria in auction games with continuous type- and action-space, i.e. sealed bid single-item auctions. We leverage the potential of GPUs to parallelize computations and find that Neural Self-Play scales well with an increasing number of parameters, finding approximate Bayes-Nash Equilibria in auction settings with 10 players within 10s of minutes on a single GPU.

After demonstrating the ability of Neural Self-Play to find Bayesian Nash equilibria in sealed bid single-item auctions, we plan to consider more complex auction designs in future research, including combinatorial and sequential auctions. For these complex auctions Neural Self-Play could benefit from more advanced architectures like recurrent neural networks that provide some sort of advanced memory ability. We plan to compare the results to those of Bosshard et al. (2017) who implemented a variant of FP for combinatorial auctions with continuous type- and action-space but whose method is run-time limited for settings with more than a few players or items.

References

Michel Benaım and Morris W Hirsch. 1999. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior* 29, 1-2 (1999), 36–72.

Martin Bichler, Alok Gupta, and Wolfgang Ketter. 2010. Research commentary—designing smart markets. *Information Systems Research* 21, 4 (2010), 688–699.

Vitor Bosshard, Benedikt Bünz, Benjamin Lubin, and Sven Seuken. 2017. Computing bayes-nash equilibria in combinatorial auctions with continuous value and action spaces. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI). Melbourne, Australia.*

Michael Bowling and Manuela Veloso. 2002. Multiagent learning using a variable learning rate. *Artificial Intelligence* 136, 2 (April 2002), 215–250.

George W Brown. 1951. Iterative solution of games by fictitious play. *Activity analysis of production and allocation* 13, 1 (1951), 374–376.

C. Daskalakis, P. Goldberg, and C. Papadimitriou. 2009. The Complexity of Computing a Nash Equilibrium. *SIAM J. Comput.* 39, 1 (Jan. 2009), 195–259.

Drew Fudenberg and David M Kreps. 1993. Learning mixed equilibria. *Games and Economic Behavior* 5, 3 (1993), 320–367.

Drew Fudenberg and David K. Levine. 1999. *The theory of learning in games* (2. ed.). MIT Press series on economic learning and social evolution, Vol. 2. MIT Press, Cambridge.

James S Jordan et al. 1993. Three problems in learning mixed-strategy Nash equilibria. *Games and Economic Behavior* 5, 3 (1993), 368–386.

Günter Klambauer, Thomas Unterthiner, Andreas Mayr, and Sepp Hochreiter. 2017. Self-Normalizing Neural Networks. *arXiv preprint, arXiv:1706.02515* (June 2017).

Gian-Marco Kokott, Martin Bichler, and Per Paulsen. 2019. The beauty of Dutch: Ex-post split-award auctions in procurement markets with diseconomies of scale. *European Journal of Operational Research* 278, 1 (2019), 202–210.

Flavio M Menezes and Paulo Klinger Monteiro. 2005. *An introduction to auction theory.* OUP Oxford.

John H Nachbar. 1990. “Evolutionary” selection dynamics in games: Convergence and limit properties. *International journal of game theory* 19, 1 (1990), 59–89.

John F Nash et al. 1950. Equilibrium points in n-person games. *Proceedings of the national academy of sciences* 36, 1 (1950), 48–49.

Michael Plum. 1992. Characterization and computation of Nash-equilibria for auctions with incomplete information. *International Journal of Game Theory* 20, 4 (1992), 393–418.

Julia Robinson. 1951. An iterative method of solving a game. *Annals of mathematics* (1951), 296–301.

Aviad Rubinstein. 2016. Settling the complexity of computing approximate two-player Nash equilibria. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS).* IEEE, 258–265.

Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, and Ilya Sutskever. 2017. Evolution strategies as a scalable alternative to reinforcement learning. *arXiv preprint arXiv:1703.03864* (2017).

Lloyd Shapley. 1964. Some topics in two-person games. *Advances in game theory* 52 (1964), 1–29.

Satinder Singh, Michael Kearns, and Yishay Mansour. 2000. Nash Convergence of Gradient Dynamics in Iterated General-Sum Games. *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI2000)* (June 2000).

6 Equilibrium Learning in Combinatorial Auctions

Peer-Reviewed Workshop Paper

Title: Equilibrium Learning in Combinatorial Auctions: Computing Approximate Bayesian Nash Equilibria via Pseudogradient Dynamics

Authors: S. Heidekrüger, P. Sutterer, N. Kohring, M. Bichler

In: Workshop on Information Technologies and Systems 2020

Abstract: Applications of combinatorial auctions (CA) are prevalent in practice, yet their Bayes-Nash equilibria (BNE) remain poorly understood. Analytical solutions are known only for a few cases where the problem can be formulated as a tractable partial differential equation (PDE). In the general case, finding BNE is known to be computationally hard. Previous work on numerical computation of BNE in auctions has relied either on solving such PDEs explicitly, calculating pointwise best-responses in strategy space, or iteratively solving subgames with restricted strategy spaces. We present a generic yet scalable method, representing strategies as neural networks and applying updates via gradient dynamics in self-play. Most CAs are ex-post nondifferentiable, so gradients are unavailable or misleading, and we instead rely on suitable pseudogradient estimates. We observe fast and robust convergence to approximate BNE in a wide variety of CA games and give a sufficient condition for convergence.

Equilibrium Learning in Combinatorial Auctions: Computing Approximate Bayesian Nash Equilibria via Pseudogradient Dynamics

Stefan Heidekrüger, Paul Sutterer, Nils Kohring, Martin Bichler
Department of Informatics, Technical University of Munich
stefan.heidekrueger,paul.sutterer,kohring,bichler@in.tum.de

November 27, 2020

Abstract

Applications of combinatorial auctions (CA) are prevalent in practice, yet their Bayesian Nash equilibria (BNE) remain poorly understood. Analytical solutions are known only for a few cases where the problem can be formulated as a tractable partial differential equation (PDE). In the general case, finding BNE is known to be computationally hard. Previous work on numerical computation of BNE in auctions has relied either on solving such PDEs explicitly, calculating pointwise best-responses in strategy space, or iteratively solving subgames with restricted strategy spaces. We present a generic yet scalable method, representing strategies as neural networks and applying updates via gradient dynamics in self-play. Most CAs are ex-post nondifferentiable, so gradients are unavailable or misleading, and we instead rely on suitable pseudogradient estimates. We observe fast and robust convergence to approximate BNE in a wide variety of CA games and give a sufficient condition for convergence.

1 Introduction

Auctions are widely used in private sectors, e.g. advertising and procurement, as well as in the public sector for the allocation of social goods, e.g. for spectrum sales and for the extension of renewable energy sources. [8, 24, 3, 32]. Auction markets inherently involve incomplete information about competitors and strategic behavior of market participants. Understanding decision making in such markets is an important line of research in information systems. Auctions are typically modeled as Bayesian games and one is particularly interested in the equilibria of such games to potentially enable welfare maximizing allocations.

It is well-known that equilibrium computation is hard. Finding Nash equilibria is known to be PPAD-complete even for normal-form games, which assume complete information and finite action spaces, and in which a Nash equilibrium is guaranteed to exist [16]. In auctions, however, these assumptions no longer hold. In auction games modeled as Bayesian games with continuous type and action spaces, agents' values are drawn from some continuous prior value distribution and their actions are typically described as a continuous bid function of these valuations. For auctions of a single item, the landmark results by Vickrey [31] have enabled a deep understanding of common auction formats. For multi-item auctions and more specifically for *combinatorial auctions*, in which players bid on *bundles* of multiple items simultaneously, there has been little progress. While the complexity of computing Bayes-Nash equilibria (BNE) is not well understood, Cai and Papadimitriou [14] show that BNE computation for a specific combinatorial auction is already (at least) PP-hard. Furthermore, finding an ϵ-approximation to a BNE is still NP-hard. Explicit solutions exist for very few specific environments, but in general, we neither know whether a BNE exists nor do we have a solution theory. Combinatorial auctions are widely used in the field [8, 15], thus understanding their equilibria is paramount, and access to scalable numerical methods for computing or approximating BNE can have a significant impact on the possibility to enable social welfare maximizing allocations.

Equilibrium learning in games differs from most learning tasks in that it suffers from the *nonstationarity problem*: Each player's objective depends on other agents' actions. Prior literature on equilibrium learning primarily focuses on complete-information games. In contrast, we focus on Bayes-Nash equilibria in auction games with continuous action space and continuous prior value distributions. The literature on equilibrium computation for these games is in its infancy and largely relies on best-response computations. In this paper, we propose Neural Pseudogradient Ascent (NPGA) as an equilibrium learning method that uses gradient dynamics. While learning based on gradient dynamics has been used in complete-information games, this is not the case for Bayesian auction games: First, the underlying problem is equivalent to an infinite-dimensional variational inequality, for which we do not know an exact solution method. Second, the (ex-post) payoff function of auction games is non-differentiable. Finally, gradient dynamics have been known to converge to Nash equilibria only in restricted games, even under complete information.

NPGA relies on self-play with neural networks, uses evolutionary strategies to compute gradients, and can exploit GPU hardware acceleration to massively parallelize the computations. In contrast to some previous work on numerical BNE computation, NPGA does not require any setting-specific subprocedures or information beyond evaluating auction outcomes themselves, and it can thus be applied to arbitrary Bayesian games. We discuss a sufficient condition for convergence of NPGA to a Bayes-Nash equilibrium and provide extensive experimental results on combinatorial auctions, which pose a benchmark problem in algorithmic game theory. Interestingly, we observe convergence of NPGA to approximate BNE in a wide range of small- and medium-sized combinatorial auction environments and recover the analytical Bayes-Nash equilibrium where it is known.

The development of NPGA as a numerical tool based on machine learning, leveraging GPU hardware, and the application to a highly relevant economical problem that can lead to significant increase in social welfare, contributes to the overall theme of the workshop: Multi-method Design of Intelligent Systems for Social Good.

The remainder of this paper is structured as follows: In Section 2, we formally introduce the model and the problem. Section 3 examines related work, both in the problem domain of combinatorial auctions and related to our methodology. Next, we introduce and discuss NPGA in Section 4, before applying it to a suite of previously studied combinatorial auctions in Section 5. Finally, we summarize our findings and outline future research directions.

2 Problem statement

Bayesian games and combinatorial auctions. A *Bayesian game* or *incomplete information game* is described by a quintuple $G = (\mathcal{I}, \mathcal{A}, \mathcal{V}, F, u)$. Here $\mathcal{I} = \{1, \ldots, n\}$ describes the set of agents participating in the game. Throughout this paper, we denote by the index $-i$ a profile of types, actions or strategies for all agents but agent i. $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ is the set of possible action profiles, with $\mathcal{A}_i$ being the set of actions available to agent $i \in \mathcal{I}$. $\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_n$ is the set of *type profiles.* $F : \mathcal{V} \to [0, 1]$ defines a common prior probability distribution over type profiles that is assumed to be common knowledge among all agents in the game. For any dependent random variable X, we denote its cumulative distribution function by F_X and its probability density function by f_X. For example, F_{v_i} denotes the marginal distribution of agent i's type. At the beginning of the game, nature draws a type profile $v \sim F$ and each agent i is informed of their own type $v_i \in \mathcal{V}_i$ only, thus the type constitutes private information based on which each agent chooses their action $b_i \in \mathcal{A}_i$. Each agent's *ex-post* utility function is then determined by $u_i : \mathcal{A} \times \mathcal{V}_i \to \mathbb{R}$, i.e. the agent's utility depends on all agents' actions but only on their own type. Agents aim to maximize their individual utility or *payoff* u_i.

In this paper, we consider *sealed-bid combinatorial auctions* (CA) on $\mathcal{M} = \{1, \ldots, m\}$ items. In such an auction, each agent, or bidder in this context, is allocated a bundle $k_i \in \mathcal{K} = 2^{\mathcal{M}}$ of items. Each agent's types $v_i \in \mathcal{V}_i$ are given by a vector of *private valuations* over bundles, i.e. $v_i = (v_i(k))_{k \in \mathcal{K}}$. Bidders then submit actions, called *bids* b_i, according to some bid-language: In the general case, where bidders might be interested in any combination of items, bids are in $\mathcal{A}_i \subseteq \mathbb{R}_+^{|\mathcal{K}|}$, i.e. each player must submit 2^m bids. In practice this is prohibitive, and one commonly studies settings where valuations exhibit some structure that allows reducing the dimensionality of the type space as well as the bid language. The settings we study in Section 5 have type and action spaces $\mathbb{R}_+$ or $\mathbb{R}_+^2$.

CAs are modeled as Bayesian games: First, nature draws a *valuation profile* $v \in \mathcal{V}$ and each bidder i observes their own type v_i. Then, bidders submit bids $b_i = \beta_i(v_i)$ chosen according to some *strategy* or *bid function* $\beta_i : \mathcal{V}_i \to \mathcal{A}_i$ that maps individual valuations to a probability distribution over possible actions.[1] We denote by $\Sigma_i \subseteq \mathcal{A}_i^{\mathcal{V}_i}$ the resulting strategy space of bidder i and by $\Sigma \equiv \prod_i \Sigma_i$ the space of possible joint strategies. Note that even for deterministic strategies, the spaces Σ_i are infinite-dimensional unless $\mathcal{V}_i$ are finite. The auctioneer collects these bids, applies some *auction mechanism* that determines (a) an allocation $x \in \mathcal{K}^n$; each bidder i receives a (possibly empty) bundle $x_i \in \mathcal{K}$, s.t. the union of these bundles is disjoint, $\dot{\bigcup}_i x_i \subseteq \mathcal{K}$, i.e. each item $m \in \mathcal{M}$ is allocated to at most one bidder, and (b) payments $p \in \mathbb{R}^n$ that the agents have to pay to the auctioneer. For brevity, we will restrict ourselves to bidders with *quasi-linear* utility functions[2] given by $u_i : \mathcal{V}_i \times \mathcal{A} \to \mathbb{R}$,

$$u_i(v_i, b_i, b_{-i}) = v_i(x_i) - p_i. \tag{1}$$

Throughout this paper, we will differentiate between the *ex-ante* state of the game, where players know only the priors F, the *ex-interim* state, where players additionally know their

[1] *Mixed* strategies that randomize over actions would also be possible, but we restrict ourselves to *pure* or *deterministic* strategies that choose a specific action with certainty, as most work in auction theory focuses on pure-strategy Bayesian Nash equilibria.

[2] This corresponds to risk-neutral bidders, but our method is applicable to arbitrary risk-profiles.

own valuation $v_i \sim F_i$, and the *ex-post* state, where all actions have been played and $u_i(v, b)$ can be observed.

Equilibria in Bayesian games. In non-cooperative game theory, Nash equilibria (NE) are the central equilibrium solution concept. A set of bids b^* is a pure-strategy NE of the complete-information game $G = (\mathcal{I}, \mathcal{A}, u)$ if $u_i(b_i^*, b_{-i}^*) \geq u_i(b_i, b_{-i}^*)$ for all $b_i \in \mathcal{A}_i$ and all $i \in \mathcal{I}$. In a NE no agent has an incentive to deviate unilaterally, given the equilibrium strategy of all other agents. Bayesian-Nash equilibria extend the standard notion of NE to incomplete-information games, calculating the expected utility $\overline{u}$ over the conditional distribution of opponent valuations v_{-i}. For a valuation $v_i \in \mathcal{V}_i$, action $b_i \in \mathcal{A}_i$ and fixed opponent strategies $\beta_{-i} \in \Sigma_{-i}$, we denote the *ex-interim utility* of bidder i by

$$\overline{u}_i(v_i, b_i, \beta_{-i}) \equiv \mathbb{E}_{v_{-i}|v_i}\left[u_i\left(v_i, b_i, \beta_{-i}(v_{-i})\right)\right]. \tag{2}$$

We also denote the *ex-interim utility loss* of action b_i incurred by not playing the best response action, given v_i and β_{-i}, by

$$\overline{\ell}_i(b_i; v_i, \beta_{-i}) = \sup_{b_i' \in \mathcal{A}_i} \overline{u}_i(v_i, b_i', \beta_{-i}) - \overline{u}_i(v_i, b_i, \beta_{-i}). \tag{3}$$

Note that $\overline{\ell}_i$ can generally not be observed in online-settings because it requires knowledge of a best-response.

An *ϵ-Bayes-Nash Equilibrium (ϵ-BNE)* is a strategy profile $\beta^* \in \Sigma$ such that no agent can improve her own ex-interim expected utility by more than $\epsilon \geq 0$ by deviating from the common strategy profile. Thus, in an ϵ-BNE, we have:

$$\overline{\ell}_i\left(b_i; v_i, \beta_{-i}^*\right) \leq \epsilon \quad \text{for all } i \in \mathcal{I}, v_i \in \mathcal{V}_i \text{ and } b_i \in \mathcal{A}_i. \tag{4}$$

A 0-BNE is simply called BNE. Thus, in a BNE, every bidder's strategy maximizes her expected ex-interim utility given opponent strategies for every possible type $v \in \mathcal{V}$. While BNE are commonly defined at the *ex-interim* stage of the game, we also consider *ex-ante* Bayesian equilibria as strategy profiles that concurrently maximize each player's *ex-ante* expected utility $\tilde{u}$. We define $\tilde{u}$ and the *ex-ante utility losses* $\tilde{\ell}$ of a strategy profile $\beta \in \Sigma$ by

$$\tilde{u}_i(\beta_i, \beta_{-i}) \equiv \mathbb{E}_v\left[u_i(v_i, \beta_i(v_i), \beta_{-i}(v_{-i}))\right] \overset{\text{eq. } 2}{=} \mathbb{E}_{v_i \sim F_{v_i}}\left[\overline{u}_i(v_i, b_i, \beta_{-i})\right], \tag{5}$$

$$\tilde{\ell}_i(\beta_i, \beta_{-i}) \equiv \sup_{\beta_i' \in \Sigma_i} \tilde{u}_i(\beta_i', \beta_{-i}) - \tilde{u}_i(\beta_i, \beta_{-i}). \tag{6}$$

Then an ex-ante BNE β^* can be characterized by the equations $\tilde{\ell}_i(\beta_i^*, \beta_{-i}^*) = 0$ for all $i \in \mathcal{I}$. Clearly, every ex-interim BNE also constitutes an ex-ante equilibrium. The reverse holds only almost surely.

3 Related work

Gradient dynamics in games have been studied in evolutionary game theory and multiagent learning. While earlier work considered mixed strategies over normal-form games [33, 12, 11, 13], more recently, motivated by the emergence of Generative Adversarial Networks, there has been a focus on (complete-information) games with continuous action spaces and smooth utility functions [23, 22, 6, 28]. A result found for many of the studied settings and algorithms is that gradient-based learning rules do not necessarily converge to Nash equilibria and may exhibit cycling behavior, but often achieve no-regret properties and thus converge to Coarse Correlated equilibria (CCE), a solution concept weaker than Nash equilibria. An analogous result exists for finite-type Bayesian games, where no-regret learners are guaranteed to converge to a Bayesian CCE [19]. In the present paper, we study equilibrium learning via gradient dynamics in *continuous-type* Bayesian games, specifically auctions, where they have not been investigated previously to our knowledge.

Equilibrium computation in auctions. Earlier approaches to find equilibria in auctions were usually setting specific and relied on reformulating Equation 4 as a differential equation (where possible), then solving this equation analytically or numerically [31, 21, 4]. Armantier et al. [2] introduced a BNE-computation method that is based on expressing the Bayesian game as the limit of a sequence complete-information games. They show that the sequence of Nash equilibria in the restricted games converges to a BNE of the original game. While this result holds for any Bayesian game, setting-specific information is required to generate and solve the restricted games. Rabinovich et al. [26] study best-response dynamics on mixed strategies in auctions with finite action spaces. Most recently, Bosshard et al. [9, 10] proposed a method to find BNE in combinatorial auctions that relies on smoothed best-response dynamics and is applicable to any Bayesian game. The method explicitly computes point-wise best-responses in a fine-grained linearization of the strategy space via sophisticated Monte-Carlo integration. With NPGA, we introduce a method that does not require setting-specific information while avoiding explicit best-response computations to remain scalable.

4 Pseudogradient dynamics in auction games

Next, we present our method for equilibrium computation in auctions. On a high level, we propose following the gradient dynamics of the game via simultaneous gradient ascent of all bidders. As we will see, however, computing the gradients themselves is not straightforward in the auction setting and we will need some modifications to established gradient dynamics methods such as [33, 29]. For now, assume that players observe a gradient-oracle $\nabla_{\beta_i}\tilde{u}_i(\beta_i, \beta_{-i})$ with respect to the current strategy profile β^t in each iteration. Then the learning rule proposes that players perform a projected gradient update:

$$\beta_i^t \equiv \mathcal{P}_{\Sigma_i}\left(\beta_i^{t-1} + \Delta_i^t\right) \quad \text{with} \quad \Delta_i^t \propto \nabla_{\beta_i}\tilde{u}_i(\beta_i, \beta_{-i}), \tag{7}$$

where $\mathcal{P}_{\Sigma_i}(\,\cdot\,)$ is the projection onto the set of feasible strategies for agent i. Several things should be noted about Equation 7: First, we consider the gradient dynamics of the *ex-ante* utility $\tilde{u}$, rather than ex-interim or ex-post utilities. The goal of an individual update step is thus to marginally improve the expected utility of player i across all possible joint valuations $v \sim F$. This perspective ultimately considers low-probability events less important than high-probability events, which is in contrast to some other methods, which explicitly aim to optimize *all* ex-interim states [10]. Second, to compute the gradient oracle $\nabla_\beta \tilde{u}$ in self-play, we rely on access to other players strategies, but evaluating each player's policy relies only on their own valuation. We thus follow the centralized-training, decentralized-execution framework common in multi-agent learning. Third, $\beta_i \in \Sigma_i$ are functions in an infinite-dimensional function space, so the gradient $\nabla_{\beta_i}\tilde{u}_i$ is itself a *functional* derivative. In our ex-ante perspective, we thus consider this to be the Gateaux derivative over the Hilbert space

Σ_i, equipped with the inner product $\langle \psi, \beta_i \rangle = \mathbb{E}_{v\sim F}\left[\psi(v)^T \beta_i(v)\right]$ (which, in turn, defines the projection in Equation 7 as $\mathcal{P}_{\Sigma_i}(\beta) \equiv \arg\min_{\sigma\in\Sigma_i} \langle \sigma - \beta, \sigma - \beta \rangle$).

Policy networks. To implement this derivative in practice, we represent each bidder's strategy by a *policy network* $\beta_i(v_i) \equiv \pi_i(v_i;\theta_i)$ specified by a neural network architecture and a corresponding parameter vector $\theta_i \in \mathbb{R}^{d_i}$. Importantly, given a suitable network architecture, one can ensure that all θ_i correspond to feasible β_i, thus making the projection in the update step obsolete. In the empirical part of this study, we restrict ourselves to fully-connected feed-forward neural networks with ReLU activations in the output layer, which ensure nonnegative bids—the only feasibility constraint in the auctions we study. In any case, $d_i \in \mathbb{N}$ is finite and we thus transform the problem of choosing an infinite-dimensional strategy into choosing a finite-dimensional parameter vector θ_i.

Policy pseudogradients. The *deterministic policy gradient theorem* [29] gives an established way to compute the payoff gradient with respect to the parameters θ:

$$\nabla_{\theta_i} \tilde{u}_i(\pi_i(\,\cdot\,;\theta_i), \beta_{-i}) = \mathbb{E}_{v\sim F}\left[\nabla_{\theta_i}\pi(v_i;\theta_i)\nabla_{b_i} u_i(v_i; b_i, \beta_{-i}(v_{-i}))|_{b_i=\pi_i(v_i;\theta_i)}\right]. \tag{8}$$

However, the regularity conditions required by the theorem are commonly violated in combinatorial auctions. In particular, due to the discrete nature of the allocations x, the ex-post utilities $u_i(v_i, b_i, b_{-i})$ are usually discontinuous—and thus neither differentiable nor subdifferentiable in b_i. While this nondifferentiability does not extend to $\tilde{u}$, it nevertheless renders the policy gradient formula in Equation 8 inapplicable. Although the set of discontinuities is a v-nullset in practice, one can show that even on the differentiable intervals of $u_i(v_i,\,\cdot\,, b_{-i})$, its true gradient provides systematically misleading signals:

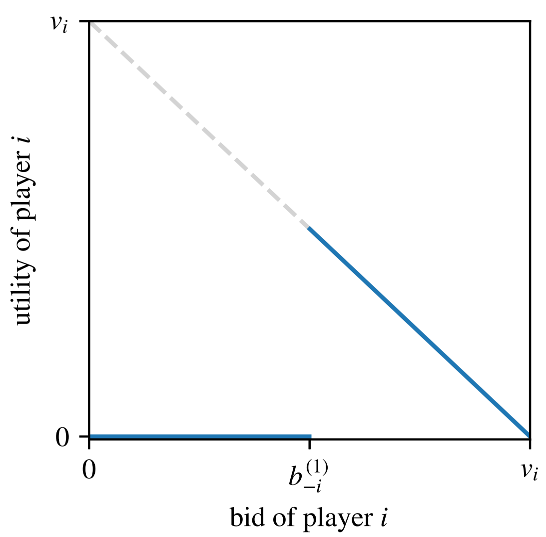


Figure 1: Utility function $u_i(b_i)$ in single-item First-Price Sealed-Bid auction for fixed opponent bids b_{-i} with highest opponent bid $b_{-i}^{(1)}$.

Consider a first-price sealed-bid auction in which winning bidders pay their bid amount b_i. The utility graph is separated into two sections (see Figure 1): (a) Bidding lower than the highest opposing bid leads to a zero payoff and thus no learning feedback, $\nabla_{b_i} u_i = 0$, and (b) winning with a high enough bid leads to a learning feedback to decrease the bid, $\nabla_{b_i} u_i = -1$. In short, backpropagation will thus lead to a steady decrease of bids in every iteration, until all players bid constant zero for any valuation.

To alleviate this, we follow the approach of [27] and estimate the policy gradient using a finite difference approach based on evolutionary strategies. To calculate $\nabla_\theta \tilde{u}$, we perturb the parameter vector P times, $\theta_{i;p} \equiv \theta_i + \varepsilon_p$, using zero-mean Gaussian noise $\varepsilon_p \sim \mathcal{N}(0, \sigma^2)$, where

σ is a hyperparameter. We then calculate each perturbation's *fitness*, $\varphi_p \equiv \tilde{u}_i(\pi_i(v_i;\theta_{i;p}), \beta_{-i})$, via Monte-Carlo integration, and estimate the gradients as the fitness-weighted perturbation noise $\nabla_\theta^{ES} \equiv \frac{1}{\sigma^2 P}\sum_p \varphi_p \varepsilon_p$. Salimans et al. [27] motivated the approach for its parallelization potential, but here we exploit its property that it gives an unbiased estimator of $\nabla_\theta \tilde{u}$ even when $\nabla_b u$ is not well-defined.

The complete pseudocode of NPGA, is described in Algorithm 1. To summarize, NPGA "implements" Equation (7) via ES-pseudogradients and a neural network parametrization of strategy functions which renders the projection step unnecessary:

$$\beta_i^t \equiv \pi_i(\,\cdot\,;\theta_i^t) \quad \text{with} \quad \theta_i^t \equiv \theta_i^{t-1} + \Delta_i^t \quad \text{where} \quad \Delta_i^t \propto \nabla_{\theta_i^t}^{ES}. \tag{9}$$

Algorithm 1: Neural Pseudogradient Ascent using Evolutionary Strategy gradients

Input: players $i \in \mathcal{I}$; ex-post utility functions u_i, initial policy networks $\beta_i^0 \equiv \pi_i(\,\cdot\,,\theta_i^0)$, with initial parameters $\theta_i^0 \in \mathbb{R}^{d_i}$; ES population size P; ES noise standard deviation σ; batch size K

for $t := 1, 2, \ldots$ **do**

Sample a batch $(v_k)_{k=1..K}$ of valuation profiles, with $v_k \sim F$

Calculate joint ex-ante utility in current strategy profile:

$$\tilde{u}^{t-1} := \frac{1}{K}\sum_k u\left(\beta^{t-1}(v_k)\right)$$

for each player $i \in \mathcal{I}$ **do**

Sample P perturbations of player i's current policy parameters:

$$\theta_{i;p} := \theta_i^{t-1} + \varepsilon_p \text{ with } \varepsilon_p \sim \mathcal{N}(0, \sigma^2 I) \text{ i.i.d. for } p \in \{1, \ldots, P\}$$

For each p, evaluate the fitness of the perturbation by playing against current opponents:

$$\varphi_p := \frac{1}{K}\sum_k u_i\left(\pi_i(v_{k,i}, \theta_{i;p}), \beta_{-i}^{t-1}(v_{k,-i})\right) - \tilde{u}_i^{t-1}$$

Calculate ES pseudogradient as fitness-weighted perturbation noise:

$$\nabla^{ES}\tilde{u}_i^{t-1} := \frac{1}{\sigma^2 P}\sum_p \varphi_p \varepsilon_p$$

Perform a (generalized) gradient update step on the current policy:

$$\Delta\theta_i^t \propto \nabla^{ES}\tilde{u}_i^{t-1}, \quad \theta_i^t := \theta_i^{t-1} + \Delta\theta_i^t, \quad \beta_i^t := \pi_i(\,\cdot\,,\theta_i^t)$$

end

end

Vectorizing auction evaluations. Remarkably, the only information about the game G needed in the computation of this learning rule is the evaluation of $\tilde{u} = \mathbb{E}_{v\sim F}[u]$ for a given strategy profile. Given a vectorized implementation of the joint ex-post utility function u, estimating $\tilde{u}$ via Monte-Carlo integration over $\mathcal{V}$ is suitable to parallel execution on hardware accelerators such as GPUs or TPUs. To this end, we built custom vectorized implementations of many common auction mechanisms using the PyTorch framework [25], allowing us to perform this Monte-Carlo estimation multiple orders of magnitude faster compared to previous numerical work on auctions. For moderately sized auction games, like those commonly studied in the literature, allocations x can be computed in a vectorized fashion via full enumeration of feasible allocations.[3] Common payment rules either have inherently vectorizable closed-form formulation (e.g. first-price auctions) or can be reformulated as the solution of a constrained linear or quadratic program (e.g. the Vickrey-Clarke-Groves (VCG) mechanism or core-selecting pricing rules [17]). To solve a large batch of the latter in parallel, we leverage a custom vectorized implementation of interior-point methods. (A similar approach has previously been used by [1].)

A convergence criterion. As discussed in Section 3, gradient dynamics do not converge to Nash equilibria in general. For differentiable, finite-dimensional, complete-information games, Mertikopoulos and Zhou [23] show that monotonicity of the payoff gradient is a sufficient condition for convergence of gradient dynamics to a unique Nash equilibrium as it leads to strict concavity of the game. Ui [30] shows an analogous result for ex-post differentiable Bayesian games, in which payoff-monotonicity guarantees the existence of a unique BNE. However, the result does not directly apply to auctions due to their ex-post nondifferentiability. Instead, we give a slightly less restrictive criterion based on ex-interim payoff monotonicity that ensures convergence of gradient dynamics and whose formulation is compatible with auction games.

Definition 1 (Ex-interim payoff monotonicity)**.** *Let $G = (\mathcal{I}, \mathcal{A}, \mathcal{V}, F, u)$ be a Bayesian game, such that the individual ex-interim utilities are continuously differentiable in b_i and the gradients are bounded by $\|\nabla \overline{u}_i(v_i, b_i, \beta_{-i})\| \leq Z$. G is called* strictly (ex-interim) payoff-monotone*, if for all $i \in \mathcal{I}$, $\beta_{-i} \in \Sigma_{-i}$, $a_i, b_i \in \mathcal{A}_i$ and almost everywhere $v_i \in \mathcal{V}_i$ the following holds:*

$$\langle \nabla_{a_i} \overline{u}_i(v_i, a_i, \beta_{-i}) - \nabla_{b_i} \overline{u}_i(v_i, b_i, \beta_{-i}), a_i - b_i \rangle \ < \ 0. \tag{10}$$

While analytical verification of this criterion is elusive, except in special settings, it can (approximately) be checked numerically. We observe it in all settings studied below.

Here, we give a convergence result under ex-interim monotonicity. For convergence analysis, we will also rely on certain properties of "appropriate" neural network architectures, most importantly preservation of concavity.

Definition 2 (NPGA Policy Network)**.** *An* NPGA Policy Network *is a neural network $\pi_i : \mathcal{V}_i \times \Theta_i \to \mathcal{A}_i$ with $\dim \Theta_i = d_i$ and the following properties:*

1. *π_i is a* convex neural network *in its parameters: For any convex objective function $g : \Sigma_i \to \mathbb{R}$, the map $\theta_i \mapsto g(\pi_i(\,\cdot\,, \theta_i))$ is convex.*

2. *π_i* universally approximates Σ_i*: There exists a $\delta > 0$, s.t. for all $\beta_i \in \Sigma_i$ there is a parameter vector $\theta_i \in \Theta_i$ with $\mathbb{E}_{v_i}\left[\|\beta_i(v_i) - \pi_i(v_i, \theta_i)\|\right] \leq \delta$.*

3. *π_i is Lipschitz-continuous in its parameters in the sense that there's some $L > 0$ such that for all $\theta_i, \theta_i' \in \Theta_i$: $\mathbb{E}_{v_i}\left[\|\pi_i(v_i, \theta_i) - \pi_i(v_i, \theta_i')\|\right] \leq L\|\theta_i - \theta_i'\|$.*

[3] We note that, with current hardware, this approach remains intractable for some larger auctions that are applied in the real world.

Neural networks that are employed in practice (and our empirical analysis) generally do not comply with Definition 2, but such networks have been shown to exist. For example Bach [5] studies wide single-hidden layer networks with ReLU activations, in which only the output-layer weights are being trained.

Proposition 1. *(simplified) Let $G = (\mathcal{I}, \mathcal{A}, \mathcal{V}, F, u)$ be a Bayesian game such that the ex-interim payoff-gradients exist and fulfill strict ex-interim payoff monotonicity. Then, with an NN architecture as in Definition 2 and appropriate update step sizes, NPGA converges to an ex-ante ϵ-BNE of G.*

While existence and uniqueness of BNE are generally unknown, Proposition 1 guarantees efficient computability in a wide range of settings, some of which we explore in the next section. Nonetheless, it is important to note that there may be auction settings where payoff-monotonicity does not hold.

An abridged proof of Proposition 1 is given in Appendix A. Due to space constraints, we omit some of the more technical derivations from the proof and focus on the instructive parts of the argument.

5 Results

We evaluate NPGA on a benchmark problem in algorithmic game theory: local-global auctions. Global bidders' priors allow them to draw higher valuations, so local bidders need to coordinate to outbid the global bidder. We consider bidders with independent and with correlated, uniform priors $F_{v_i} = \mathcal{U}(0, \overline{v}_i)$ with $\overline{v}_l = 1, \overline{v}_g = 2$ for local (l) and global (g) bidders. The standard LLG setting (Section 5.1) has been widely studied [4, 18] for theoretical insight. The larger LLLLGG setting (Section 5.2) was proposed by [10]. To our knowledge, it is the most complex environment in which BNE have been computed. We analyze NPGA in six settings where an analytical BNE is known and three settings where it is not.

We use common hyperparameters across all settings: Fully connected neural networks with two hidden layers and SeLU activations [20] in the hidden layers, $P = 64$, Adam optimizer steps with a learning rate of 0.001, and other hyperparameters as suggested in the respective original paper. To avoid degenerate initializations of θ (e.g. where one or more bidders bid constant zero), we perform supervised pre-training to the *truthful strategy* $\beta_i(v_i) = v_i$. All experiments were performed on a single Nvidia GeForce 2080Ti and batch sizes in Monte-Carlo sampling were chosen to maximize GPU-RAM utilization: A learning batch size of $K = 2^{18}$; An evaluation batch size (for $\ell^*, ||\beta - \beta^*||$) of $H = 2^{22}$; A secondary evaluation batch size $H = 2^{12}$ and grid size $W = 2^{10}$ (for $\hat{\ell}, \hat{\epsilon}$).

Each experiment is repeated ten times over 5,000 iterations each. Only the parameters for LLLLGG with nearest-vcg differ due to its complexity, i.e. nearest-vcg requires solving a linear- and a subsequent quadratic optimization problem for each individual auction [17]. Because NPGA solves many thousand auctions at once, we implemented a solver that allows to solve batches of quadratic optimization problems on the GPU. Nonetheless, we had to adjust the parameters to: $P = 32$; $K = 2^{14}$; $H = 2^7$ and $W = 2^8$; Experiments are repeated two times with 1,000 iterations each.

Table 1: Results of NPGA in the LLG auctions with independent and correlated valuations. [1]Estimating the utility loss with correlated priors is not straightforward at all but will be part of future research and is not needed here since ℓ^* is known.

priors	**payments**	**bidder**	ℓ^*	$\lVert\beta-\beta^*\rVert$	$\hat{\ell}$	$\hat{\epsilon}$	**conv.** iters	**sec** per iter
independent	**nearest-vcg**	**locals**	0.0001	0.0050	0.0002	0.0009	320	0.84
		global	0.0000	0.0269	0.0000	0.0001		
	nearest-bid	**locals**	-0.0002	0.0073	0.0003	0.0013	540	0.79
		global	0.0000	0.0424	0.0000	0.0001		
	nearest-zero	**locals**	-0.0001	0.0078	0.0002	0.0019	630	0.79
		global	0.0000	0.0088	0.0000	0.0001		
	FPSB	**locals**	–	–	0.0009	0.0031	1,140	0.65
		global	–	–	0.0016	0.0064		
correlated	**nearest-vcg**	**locals**	-0.0001	0.0042	[1]–		–	0.80
		global	0.0000	0.0305	–		–	
	nearest-bid	**locals**	0.0003	0.0064	–		–	0.83
		global	0.0000	0.0498	–		–	
	nearest-zero	**locals**	0.0001	0.0059	–		–	0.81
		global	0.0000	0.0072	–		–	

5.1 The LLG setting

The LLG setting includes two local bidders and one global bidder that bid on $m = 2$ items. Local bidders $i = 1, 2$ are each interested in the bundle $\{i\}$, while the global bidder wants the package $\{1, 2\}$ of both items. Each bidder submits a bid $b_i \in \mathbb{R}_+$ for their respective bundle. The setting has been extensively studied in the context of different *core-selecting* pricing rules as they are used in spectrum auctions [17, 18]. Closed-form solutions of the unique, symmetric BNE under three such rules are known in the LLG setting for independent and correlated priors: the nearest-VCG rule, the nearest-zero (or proxy) rule, and the nearest-bid rule. The interested reader is referred to [4] for details. For all three core payment rules, it has been shown that the global bidder is bidding truthfully in the BNE. The local bidders' BNE strategies differ in each payment rule. We evaluate NPGA on the three core payment rules with independent and correlated priors as well as the first-price payment rule with independent priors, for which no exact BNE is known.

Result 1. *NPGA converges to the BNE in all six settings with a known BNE after only a few hundred iterations. The benefit of deviating from the learned strategy in the first-price auction is minuscule.*

Figure 2 depicts the strategy learned by NPGA after 5,000 iterations in comparison to the analytical BNE strategy for the nearest-zero payment rule, and shows an almost perfect fit. Numerical results for all rules are presented in Table 1. For the six settings with known analytical equilibrium strategies, we measure each player's utility loss ℓ_i^* that results from unilaterally deviating from the BNE strategy profile β^* by playing the learned strategy β_i instead: $\ell^* \equiv \tilde{\ell}_i(\beta_i, \beta^*_{-i})$ (compare Equation 6.) Because of the close to perfect fit, the losses

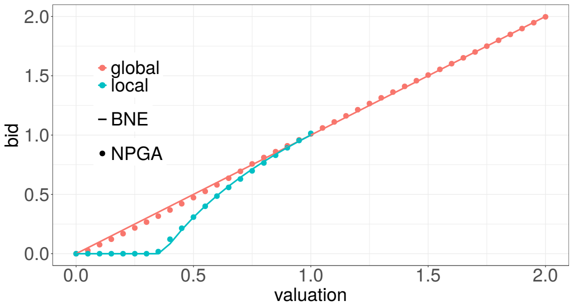


Figure 2: Learned strategies in LLG with nearest-zero core payment rule. The solid lines indicate the BNE strategies, the dots indicate the NPGA bids.

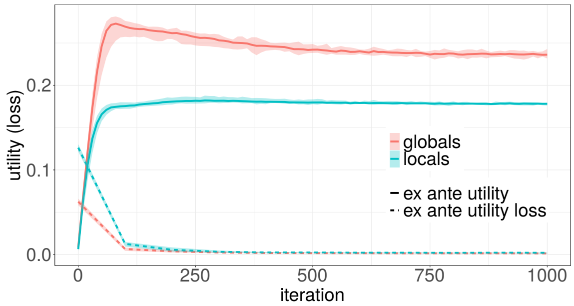


Figure 3: Ex-ante utility $\tilde{u}$ and estimated loss $\hat{\ell}$ of in NPGA self-play in the LLLLGG first-price auction. Shaded area and line show min, max, mean over 10 repetitions.

in all core payment rules are minimal, sometimes even negative. This is due to sampling noise. We also confirm the close to perfect fit by measuring the L^2-distance in strategy-space between the functions β_i and β_i^*. Here, we still observe a small deviation.

For the first price payment rule, no BNE strategy profile is known. Therefore, we estimate the potential gains of deviating from the current strategy profile $\hat{\ell}_i \approx \tilde{\ell}_i(\beta_i; \beta_{-i})$ as well as an estimator $\hat{\epsilon}$ to the "true" epsilon of β (such that β forms an ex-interim ϵ-BNE).

For the calculation of the estimators $\hat{\ell}$ and $\hat{\epsilon}$ we additionally introduce a grid of actions covering the action spaces $\mathcal{A}_i$, and consisting of W equidistant points per player. Using this grid, and a given valuation v_i and bid b_i, we then estimate the ex-interim utility loss $\bar{\ell}(v_i, b_i, \beta_{-i})$ of b_i at v_i via

$$\hat{\lambda}_i(v_i; b_i, \beta) \equiv \frac{1}{H} \max_w \sum_h u\left(v_i; b_w, \beta_{-i}(v_{h,-i})\right) - u\left(v_i; b_i, \beta_{-i}(v_{h,-i})\right). \tag{11}$$

Note that the batch H only runs across opponent valuations v_{-i}. To evaluate $\hat{\lambda}_i$ at a single v_i, we thus need $(W+1) \cdot H$ auction evaluations. We then evaluate

1. the worst-case ex-interim loss: $\hat{\epsilon} = \max_h \hat{\lambda}_i(v_{h,i}; \beta_i(v_{h,i}), \beta_{-i})$,
2. the ex-ante loss: $\hat{\ell} = \frac{1}{H} \sum_h \hat{\lambda}_i(v_{h,i}; \beta_i(v_{h,i}), \beta_{-i})$.

Estimations of $\hat{\lambda}$ can be shared for both computations, nevertheless we need $\mathcal{O}(nWH^2)$ auction evaluations to calculate these metrics (in contrast, NPGA requires $\mathcal{O}(nPK)$ evaluations, with $K < H$, $P \ll W$). This is prohibitive with large H, thus we use $W = 2^{10}, H = 2^{12}$ for the experiments and calculate $\hat{\ell}$ and $\hat{\epsilon}$ once every 100 iterations of Algorithm 1.

Both, $\hat{\ell}$ and $\hat{\epsilon}$ are small, considering the average utility of the global bidder is 0.426 and the average utility of each of the local bidders is 0.149 in the first-price auction. Note that estimating $\hat{\ell}$ and $\hat{\epsilon}$ is computationally expensive, and is not needed to run NPGA itself. In practice, these may be used as a stopping criterion to measure convergence, which we likewise report in Table 1: Once the maximum difference of the last three consecutive measurements of $\hat{\ell}$ decreases below 0.0001, we conclude that NPGA has converged. The average total computation time for all 5,000 iterations is less than 70 minutes in any LLG setting. Learning

Table 2: Results of NPGA after 5,000 (1,000) iterations in the LLLLGG first-price (nearest-vcg) auction. Results are averages over ten (two) repetitions and the standard deviation displayed in brackets. [1]Only the weaker 0.0005 criterion is met. [2]Converging criterion was met in only one of the two runs.

payments	**bidder**	$\hat{\ell}$	$\hat{\epsilon}$	**conv.** (iters)	**sec** per iter
first-price	**locals**	0.0015 (0.0003)	0.0109 (0.0025)	700[1]	0.97 (0.005)
	globals	0.0010 (0.0002)	0.0077 (0.0016)		
nearest-vcg	**locals**	0.0013 (0.0003)	0.0052 (0.0012)	600[1,2]	275.22 (0.670)
	globals	0.0011 (0.0006)	0.0098 (0.0059)		

the core payment rules with independent priors converges in less than 1,140 iterations.[4]

5.2 The LLLLGG setting

In the LLLLGG setting, four local and two global bidders compete for six items, where each bidder is interested in two (partly overlapping) bundles (containing 2 (local) or 4 (global) items each), with actions being represented as $\mathcal{A}_i = \mathbb{R}_+^2$. For this environment, no analytical BNE are known (except for the trivial VCG pricing rule, where bidding truthfully constitutes a BNE). We apply NPGA to LLLLGG with a first-price and with a nearest-vcg rule.

Result 2. *In the LLLLGG first-price and nearest-vcg auction, NPGA learns strategy profiles with an estimated ex-ante utility loss* $\hat{\ell} < 0.002$ *for both local and global bidders and both payment rules. Bidders achieve stable average utilities of 0.238 in first-price and 0.181 in nearest-vcg for each global and of 0.18, respectively 0.201, for each local bidder.*

As shown in Figure 3, applying NPGA, the bidders' utility as well as the estimated ex-ante utility losses converge fast, the latter reaching the weaker 0.0005 stopping criteria after only 700 iterations. The estimated ex-ante utility loss for both payment rules is $\hat{\ell} < 0.002$ after the complete 5,000 (1,000 in nearest-vcg) iterations, resulting in an estimated ex-ante utility loss of below 1% considering the bidders' average utilities.

6 Conclusion and future work

In this paper, we focus on equilibrium learning in Bayesian games. Gradient dynamics have been used with some success in a restricted set of complete-information games. This approach is challenging in Bayesian auction games for several reasons: these games are not differentiable, and the continuous type- and action spaces lead to challenging infinite-dimensional variational inequalities that need to be solved. We propose Neural Pseudogradient Ascent as a numerical method to compute Bayesian Nash equilibria. NPGA uses evolutionary strategies to compute

[4]The computation of estimated utility loss is not straightforward considering correlated priors since the sampling of each bidder's valuations depend on each other. Therefore, we did not compute the estimated utility loss for this setting and cannot report on the convergence criterion being met. However, the more relevant utility loss (playing against BNE) is so low that we assume convergence.

ex-ante gradients of a game that is not (sub)differentiable ex-post. In experiments, we focus on combinatorial auctions, which in general are highly relevant in practice, among others for the allocation of social goods. NPGA converges to an approximate BNE for central benchmark problems in this field. In summary, the method can provide an effective numerical tool to compute approximate BNE not only for combinatorial auctions but also for other types of Bayesian games and can help finding higher social welfare solutions in auctions.

References

[1] B. Amos and J. Z. Kolter. OptNet: Differentiable Optimization as a Layer in Neural Networks. *arXiv:1703.00443 [cs, math, stat]*, Oct. 2019.

[2] O. Armantier, J.-P. Florens, and J.-F. Richard. Approximation of Nash equilibria in Bayesian games. *Journal of Applied Econometrics*, 23(7):965–981, Nov. 2008. ISSN 08837252, 10991255. doi: 10.1002/jae.1040.

[3] I. Ashlagi, D. Monderer, and M. Tennenholtz. Simultaneous ad auctions. *Mathematics of Operations Research*, 36(1):1–13, 2011.

[4] L. M. Ausubel and O. Baranov. Core-selecting auctions with incomplete information. *International Journal of Game Theory*, July 2019. ISSN 1432-1270. doi: 10.1007/s00182-019-00691-3.

[5] F. Bach. Breaking the Curse of Dimensionality with Convex Neural Networks. *Journal of Machine Learning Research*, 18(19):1–53, 2017. ISSN 1533-7928.

[6] D. Balduzzi, S. Racaniere, J. Martens, J. Foerster, K. Tuyls, and T. Graepel. The Mechanics of n-Player Differentiable Games. *arXiv:1802.05642 [cs]; ICML*, Feb. 2018.

[7] H. H. Bauschke and P. L. Combettes. Lower Semicontinuous Convex Functions. In H. H. Bauschke and P. L. Combettes, editors, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, pages 129–141. Springer, New York, NY, 2011.

[8] M. Bichler and J. K. Goeree. *Handbook of Spectrum Auction Design.* Cambridge University Press, 2017.

[9] V. Bosshard, B. Bünz, B. Lubin, and S. Seuken. Computing Bayes-Nash Equilibria in Combinatorial Auctions with Continuous Value and Action Spaces. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 119–127, Aug. 2017.

[10] V. Bosshard, B. Bünz, B. Lubin, and S. Seuken. Computing Bayes-Nash Equilibria in Combinatorial Auctions with Verification. *Journal of Artificial Inelligence Research*, 2020.

[11] M. Bowling. Convergence and No-Regret in Multiagent Learning. In L. K. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems 17*, pages 209–216. MIT Press, 2005.

[12] M. Bowling and M. Veloso. Multiagent learning using a variable learning rate. *Artificial Intelligence*, 136(2):215–250, Apr. 2002. ISSN 0004-3702. doi: 10.1016/S0004-3702(02)00121-2.

[13] L. Busoniu, R. Babuska, and B. De Schutter. A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(2):156–172, Mar. 2008. ISSN 1094-6977. doi: 10.1109/TSMCC.2007.913919.

[14] Y. Cai and C. Papadimitriou. Simultaneous bayesian auctions and computational complexity.

In *Proceedings of the Fifteenth ACM Conference on Economics and Computation - EC '14*, pages 895–910, Palo Alto, California, USA, 2014. ACM Press. ISBN 978-1-4503-2565-3.

[15] P. Cramton, Y. Shoham, and R. Steinberg. Combinatorial Auctions. Technical Report 04mit, University of Maryland, Department of Economics - Peter Cramton, 2004.

[16] C. Daskalakis, P. Goldberg, and C. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *SIAM Journal on Computing*, 39(1):195–259, Jan. 2009. ISSN 0097-5397. doi: 10.1137/070699652.

[17] R. W. Day and P. Cramton. Quadratic Core-Selecting Payment Rules for Combinatorial Auctions. *Operations Research*, 60(3):588–603, June 2012. ISSN 0030-364X.

[18] J. K. Goeree and Y. Lien. On the impossibility of core-selecting auctions. *Theoretical Economics*, 11(1):41–52, 2016. ISSN 1555-7561. doi: 10.3982/TE1198.

[19] J. Hartline, V. Syrgkanis, and E. Tardos. No-Regret Learning in Bayesian Games. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 3061–3069. Curran Associates, Inc., 2015.

[20] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter. Self-Normalizing Neural Networks. *arXiv:1706.02515 [cs, stat]*, June 2017.

[21] V. Krishna. *Auction Theory*. Academic press, 2009.

[22] A. Letcher, D. Balduzzi, S. Racaniere, J. Martens, J. Foerster, K. Tuyls, and T. Graepel. Differentiable Game Mechanics. *arXiv:1905.04926 [cs, stat]*, May 2019.

[23] P. Mertikopoulos and Z. Zhou. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173(1):465–507, Jan. 2019. ISSN 1436-4646.

[24] P. Milgrom. *Discovering Prices: Auction Design in Markets with Complex Constraints*. Columbia University Press, 2017.

[25] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in PyTorch. In *NIPS-W*, 2017.

[26] Z. Rabinovich, V. Naroditskiy, E. H. Gerding, and N. R. Jennings. Computing pure Bayesian-Nash equilibria in games with finite actions and continuous types. *Artificial Intelligence*, 195: 106–139, Feb. 2013. ISSN 00043702. doi: 10.1016/j.artint.2012.09.007.

[27] T. Salimans, J. Ho, X. Chen, S. Sidor, and I. Sutskever. Evolution Strategies as a Scalable Alternative to Reinforcement Learning. *arXiv:1703.03864 [cs, stat]*, Mar. 2017.

[28] F. Schaefer and A. Anandkumar. Competitive Gradient Descent. *Advances in Neural Information Processing Systems 32*, pages 7625–7635, 2019.

[29] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller. Deterministic Policy Gradient Algorithms. In *International Conference on Machine Learning*, pages 387–395, Jan. 2014.

[30] T. Ui. Bayesian Nash equilibrium and variational inequalities. *Journal of Mathematical Economics*, 63:139–146, Mar. 2016. ISSN 0304-4068. doi: 10.1016/j.jmateco.2016.02.004.

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

[32] J. Winkler, M. Magosch, and M. Ragwitz. Effectiveness and efficiency of auctions for supporting

renewable electricity – what can we learn from recent experiences? *Renewable Energy*, 119:473 – 489, 2018. ISSN 0960-1481.

[33] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, pages 928–935, Washington, DC, USA, Aug. 2003. AAAI Press.

A Proof of Proposition 1

Proof (abridged). First, existence of the ex-interim gradients (required by ex-interim payoff monotonicity) implies that the ex-ante utilities $\tilde{u}_i(\beta_i, \beta_{-i}) = \mathbb{E}_{v_i}\left[\overline{u}_i(v_i, \beta(v_i), \beta_{-i}]\right)$ are Gâteaux-differentiable in the Hilbert spaces Σ_i with Gâteaux-gradients $\nabla_{\beta_i}\tilde{u}_i[\beta](v_i) = \nabla_{b_i}\overline{u}_i(v_i, b_i, \beta_{-i})|_{b_i=\beta_i(v_i)}$. (Follows from 2.54, 2.55, 17.10 of [7] and direct calculations.) With ex-interim payoff-monotonicity, we then have

$$\begin{aligned} &\langle \nabla_{\beta_i}\tilde{u}_i[\beta_i, \beta_{-i}] - \nabla_{\alpha_i}\tilde{u}_i[\alpha_i, \beta_{-i}], \beta_i - \alpha_i\rangle_{\Sigma_i} \\ = \ &\mathbb{E}_{v_i}\left[\langle \nabla_{b_i}\overline{u}_i(v_i, b_i, \beta_{-i}) - \nabla_{a_i}\overline{u}_i(v_i, a_i, \beta_{-i}),\ b_i - a_i\rangle\,|_{b_i=\beta_i(v_i), a_i=\alpha_i(v_i)}\right] \overset{eq.10}{<} \quad \mathbb{E}_{v_i}[0] = 0. \end{aligned} \tag{12}$$

I.e. the ex-ante gradients are strictly monotone operators on Σ_i. It follows that $\tilde{u}_i$ are strictly concave in β_i [see e.g. 7, Thm 17.10].

With a convex neural network as in Definition 2, the functions $\breve{u}_i(\theta_i) \equiv \tilde{u}_i(\pi_i(\,\cdot\,, \theta_i), \beta_{-i})$ are then also strictly concave in θ_i for any opponent strategies β_{-i}. We can then construct a finite-dimensional *Parameter Game*, in which all players approximate their strategies using policy networks and only choose their neural net parameters as actions: $\breve{G} \equiv (\mathcal{I}, \Theta, \breve{u})$. As this game is finite-dimensional and concave, Mertikopoulos and Zhou [23] establish that (1) it has a unique Nash equilibrium $\breve{\theta}$ and (2) the *dual averaging* algorithm converges almost surely to $\breve{\theta}$ given an unbiased and finite-variance oracle of the gradients $\nabla_{\theta_i}\breve{u}_i(\theta_i; \theta_{-i})$. We will first argue that $\breve{\theta}$ induces an approximate BNE in the original Bayesian Game G before showing that Algorithm 1 implements dual averaging in $\breve{G}$ with noisy feedback, thus finding a good approximation of $\breve{\theta}$. Let $\breve{\theta}$ thus be the Nash equilibrium of $\breve{G}$. Then for any player i, $\breve{\theta}_i$ is a best response (BR) to $\breve{\theta}_{-i}$ and $\pi_i(\,\cdot\,, \breve{\theta}_i)$ is an ex-ante BR to $\pi_{-i}(\,\cdot\,, \breve{\theta}_{-i})$ on the Bayesian Game with *restricted strategy space* of functions expressible by the neural network architecture. As we assumed universal approximation properties of π_i, however, any BR β_i^* in the *unrestricted* game G must be close in function space to $\pi_i(\,\cdot\,, \breve{\theta}_i)$, and the ex-ante utility loss incurred by not playing β_i^* instead of $\pi_i(\,\cdot\,; \breve{\theta}_i)$ is bounded: In fact, with the Lipschitz-regularity conditions on the ex-interim gradients (Definition 1) and universal approximability of π (Definition 2), one can show the following for arbitrary θ_{-i}: (We omit the details.) If $\breve{\theta}^* \in \Theta_i$ and $\beta_i^* \in \Sigma_i$ are BRs to θ_{-i} in $\breve{G}$ and G, respectively, then

$$\tilde{\ell}_i(\breve{\theta}^*; \theta_{-i}) = \tilde{u}_i(\beta_i^*, \theta_{-i}) - \tilde{u}_i(\breve{\theta}_i^*, \theta_{-i}) \leq Z\delta \tag{13}$$

where $\tilde{u}(\theta_i, \theta_{-i}) \equiv \tilde{u}_i(\pi_i(\,\cdot\,; \theta_i), \pi_{-i}(\,\cdot\,, \theta_{-i}))$. In the NE, all $\breve{\theta}_i$ are BRs, so we have $\tilde{\ell}(\breve{\theta}) \leq Z\delta$.

Finally, we show that NPGA finds a good approximation of $\breve{\theta}$. As noted in , we choose the NN architecture in such a way that Θ becomes unconstrained, i.e. any parameter $\theta_i \in \mathbb{R}^{d_i}$ is feasible, where d_i is the dimension of the network for player i. On the unconstrained action set Θ, however, dual averaging is equivalent to Online Gradient Ascent on $\breve{u}$. As deliberated above, NPGA implements Online Gradient Descent on $\tilde{u}$ using the gradient oracle ∇_θ^{ES}.

To use the convergence result of Mertikopoulos and Zhou [23] of NPGA to $\breve{\theta}$, it would remain to show that the Neural Pseudogradients $\nabla^{ES}\tilde{u}$ are finite-variance and unbiased estimators of the true gradients $\nabla_\theta \breve{u}$. This is unfortunately violated for strictly positive ES-noise variance σ^2 as used by NPGA (but asymptotically true for $\sigma \to 0$). However, for $\sigma > 0$ we can set $\breve{u}_i^\sigma \equiv \mathbb{E}_{\varepsilon\sim\mathcal{N}(0,\sigma^2 I)}\left[\breve{u}_i(\theta_i+\varepsilon,\theta_{-i})\right]$ and introduce yet another finite-dimensional game $\breve{G}^\sigma = (\mathcal{I},\Theta,\breve{u}^\sigma)$. Now, one can show (We again omit the details.) that (1) $\breve{G}^\sigma$ is, again, concave, that (2) the ES-gradients are finite-variance and unbiased estimators of $\breve{u}^\sigma$, and (3) that the loss of any parameter choice θ_i in $\breve{G}$ is bounded by that in $\breve{G}^\sigma$ via

$$\breve{\ell}_i(\theta_i,\theta_{-i}) \leq \breve{\ell}_i^\sigma(\theta_i,\theta_{-i}) + 2ZL\sqrt{d_i}\sigma. \tag{14}$$

Due to (1), $\breve{G}^\sigma$ again admits a unique NE θ^* (Theorem 2.2 of [23]), and with (2) and Definition 1 NPGA converges to θ^* almost surely for appropriate step sizes (Corollary 4.8 of [23]).

To summarize, we showed that NPGA finds a parameter profile θ^* that forms a NE of $\breve{G}^\sigma$ and which retains an-ex ante loss in G of

$$\begin{aligned}\tilde{\ell}_i(\theta^*) &= \tilde{u}_i(\beta_i^*,\theta_{-i}^*)\underbrace{-\tilde{u}_i(\breve{\theta}_i^*,\theta_{-i}^*) + \tilde{u}_i(\breve{\theta}_i^*,\theta_{-i}^*)}_{=0} - \tilde{u}_i(\theta^*)\\ &= \tilde{\ell}_i(\breve{\theta}_i^*;\theta_{-i}^*) + \breve{\ell}_i(\theta^*) \overset{(13),(14)}{\leq} Z\delta + 2ZL\sqrt{d_i}\sigma + \underbrace{\breve{\ell}_i^\sigma(\theta^*)}_{\text{loss in } \breve{G}^\sigma = 0}.\end{aligned} \tag{15}$$

Thus, setting $\epsilon \equiv Z(2L\sqrt{d}\sigma+\delta)$ where $d \equiv \max_i d_i$, NPGA converges almost surely to an ex-ante ϵ-BNE of G. $\square$

7 Discussion and Contribution

Research on auctions is manifold, addressing the topic from different angles. Some research focuses on the bidders' perspective and the problem of optimal bidding, e.g. in sequential (online ad) auctions (Zhang et al., 2014, 2015; Zhang and Wang, 2015; Perlich et al., 2012; Sutterer et al., 2019). In some combinatorial or sequential auction settings this can be formulated as a Knapsack problem (Berg et al., 2010; Chakrabarty et al., 2008). A strong game theoretical solution concept to solve Bayesian games such as auctions are Bayes-Nash equilibria (BNE). However, finding BNE is difficult, in fact at least PPAD-complete (Daskalakis et al., 2009), and little research has focused on computing BNE in auctions thus far. Only recently, the computation of (approximate) BNE in auctions has received increased attention by the research community (Bosshard et al., 2017, 2020; Heidekrüger et al., 2019, 2020b,a).

Other research focuses on the preceding design of auctions and its properties from a mechanism design perspective. The goal is to design a mechanism that implies properties such as efficiency, incentive compatibility, etc. to incentivise bidders to behave in a certain way. Typically this research focuses on multi-item and combinatorial auction settings, e.g. spectrum sales (Ausubel et al., 2006; Bichler and Goeree, 2017), fishery (Bichler et al., 2019), and renewable energy (Bichler et al., 2020; del Río, 2017).

The contribution of this thesis is threefold. Adding to the literature of bidding strategies, we first evaluate truthful bidding in display ad auctions (Sutterer et al., 2019), and second, we develop an algorithm for the computation of BNE in auctions (Heidekrüger et al., 2019, 2020b,a). Thirdly, adding to the literature of auction design, we develop an efficient alternative auction design for renewable energy sources (Bichler et al., 2020). Finally, we discuss how our research integrates into the existing literature and emphasize our contribution.

7.1 Truthful Bidding in Display Ad Auctions

With the rise of electronic markets an increasing number of auctions are conducted online. Many of them are performed in sequence, selling similar items, e.g. impressions in online display ad auctions. Bidders participating in these auctions are often limited by budgets or campaign targets. As a result, some of these sequential or combinatorial auctions can be modeled as knapsack problems.

Chakrabarty et al. (2008) study online knapsack problems with very small items compared to the knapsack capacity. Their algorithm starts with picking any items early on. As the knapsack fills it becomes more selective and only selects items that have a value higher than the competitive ratio ($^{value}/_{weight}$). In the context of online ad auctions this could be a demand-side-platform that bids on behalf of its customers, i.e. advertisers, and assigns the item to the advertiser who values it the most and has much budget left.

Zhang et al. (2014) propose a non-linear function approximation to learn optimal bidding in display ad auctions when bidders are restricted by a strict budget. In contrast to earlier work their bidding functions propose to bid on more low valued impressions instead of focusing on few high valued ones.

In Sutterer et al. (2019) (see Chapter 3), we study display ad auctions as a sequence of second-price auctions and assume for bidders to be constrained by a campaign target, i.e. a weaker notion of budgets, that models a value function with decreasing marginal valuation for additional impressions. As a result, the bidders' utility functions are quasi-linear and a theoretical offline auction, based on the Vickrey-Clarke-Groves mechanism, is incentive-compatible. The optimal offline auction serves as a benchmark for the sequential second-price online auction. We evaluate the effect of truthful bidding in online display ad auctions on efficiency and bidders' utility.

In two theoretical contributions, we show that with a typical value model for display ad auctions the outcome of a Vickrey-Clarke-Groves auction is not guaranteed to be in the core. We also show that the efficiency of an online display ad auction might be as low as 50%.

Next, we perform numerical experiments to better understand the average case efficiency and bidders' payoff when they report truthfully. We find that the welfare achieved in the online display ad auction with truthful bidders is very high, close to optimal, and the worst-case analysis is too pessimistic as an estimator for the efficiency of such auctions in the field. Furthermore, with only a few competitors bidders are able to benefit from moderate bid shading while they risk making a loss once the number of competitors increases. These results provide important information for both auctioneers and bidders in the field and contributes to the literature of bidding strategies in display ad auctions.

7.2 An Alternative Auction Design for Renewable Energy Auctions

In the course of the turnaround in global energy policy, research on the auction design of renewable energy sources has intensified. Many of the studies evaluate the overall tender designs of renewable energy sources. They identify lacking grid connections, inadequate bid bonds, low support payments, and missing penalties as the major problem that disincentives project realization (del Río and Linares, 2014; Gallachóir et al., 2009; McLean et al., 2007; Tiedemann et al., 2016; Wigand et al., 2016). Further, del Río (2017); Grashof (2013); Hauser et al. (2014); Hauser and Kochems (2014) identify and emphasize the trade-off between cost efficient support payments that meet the expansion targets and bidder diversity.

Closest to our study is the article by Anatolitis and Welisch (2017). They develop an agent-based simulation to compare the effects of a pay-as-bid and a uniform price payment rule in the context of the German onshore wind auctions. They find that prices, as well as the participation of weaker market participants, decrease each year. On average, pay-as-bid leads to lower remuneration prices than a uniform payment rule at the expense of the bidders' margin and of the efficiency.

In Bichler et al. (2020) (see Chapter 4), we propose a combinatorial auction design that allows precise steering of tendered capacities to ensure a welfare maximizing distribution (Grimm et al., 2017), while enabling project developers to explicitly communicate synergies in the form of combinatorial bids. Moreover, it is nearly strategy-proof in the large and eases the strategic bidding problem for bidders. As part of a counterfactual analysis, we compare this design to three other auction designs in terms of remuneration prices, allocation efficiency, and bidder diversity. Considering moderate synergy effects this design hardly increases the remuneration price while significantly reducing the necessity of expensive grid expansion. The results give a clear indication of the advantages over existing auction designs that ignore the location of wind power plants and their consequence on network congestion. The proposed design is not limited to the German market only but generally applicable.

With this study, we are the first to combine findings from auction design system optimal distribution of renewable energy sources. It contributes to the literature of combinatorial auction design in renewable energy sources and provides governments with a policy relevant alternative to the currently applied auction designs that can be effectively implemented.

7.3 Computing Bayes-Nash Equilibrium Strategies in Auctions

Only little research in equilibrium learning explicitly focuses on auctions. Early research in equilibrium learning of complete information and discrete state and action space (normal-form games) can be classified as *best response dynamics*. In Cournot (1838) players play a pure-strategy best-response according to the opponents' play in the previous iteration. Later, Brown (1951) propose Fictitious Play where players play a pure-strategy best-response according to the distribution of the opponents' complete past play. In its refinement, Smooth Fictitious Play, a player's play is no longer a pure-strategy but a probability distribution of possible actions. If the empirical frequencies of Fictitious or Smooth Fictitious Play converge, the limit constitutes a Nash equilibrium.

To the best of the author's knowledge, games of incomplete information, i.e. Bayesian Games, with continuous state and action space and specifically one of its most prominent applications, i.e. auctions, are not well studied thus far. Bosshard et al. (2017, 2020) were one of the first to compute approximate BNE in more complex, namely local-global combinatorial, auctions. Their approach computes point-wise best-responses in a fine-grained linearization of the strategy space via sophisticated Monte-Carlo integration. Assuming independent priors and risk-neutrality of agents, their verification method can guarantee an upper bound on the ex-interim loss in payoff, thus provably finds an ε-BNE.

In Heidekrüger et al. (2019) and Heidekrüger et al. (2020a) (see Chapter 5 and 6), we develop an algorithm that we call Neural Pseudogradient Ascent (NPGA) to learn approximate BNE. NPGA works entirely differently compared to prior best-response algorithms. It learns bid functions for all possible values (rather than point-wise) by searching the parameter space of the bid function via ex ante gradient ascent. It can be adapted to various types of Bayesian games with low re-implementation effort.

We first apply NPGA to complete information normal-form games as well as simple first-price auctions with symmetric and asymmetric bidders (Heidekrüger et al., 2019). We extend the experiments to Bayesian settings with risk-averse or risk-neutral bidders and multi-unit auctions (Heidekrüger et al., 2020b). Finally, we apply NPGA to more complex combinatorial auctions with different core payment rules and independent or correlated priors (Heidekrüger et al., 2020a).

The results of our numerical study confirm that NPGA has the same potential as Fictitious and Smooth Fictitious Play to compute Nash equilibria in normal-form games. Further, it is able to compute approximate BNE for a wide range of different auction games with continuous type and action space. In auctions with multiple BNE it is especially difficult to decide which BNE should be played, yet in our experiment NPGA always converged to the Pareto-dominant BNE. Leveraging GPU-hardware, it is possible to compute thousands of auctions in one batch and to solve even difficult combinatorial auction games on consumer hardware within hours.

The value of NPGA to the economic community can be significant. BNE are considered very effective for predicting market participants' actions. However, finding BNE is an extremely difficult task, in fact at least PPAD-complete. With NPGA we provide an algorithm that allows to efficiently learn strategies in (auction) games that often converge to the BNE. In any case, it can provide an ε-BNE estimate. This study contributes to the literature of computational techniques to find BNE and can help bidders find good bidding strategies in general auction games.

8 Conclusion

Auctions have become a prominent mechanism for the distribution of goods. The rise of electronic markets has enabled a rapid increase in their application. Today, auctions are omnipresent, e.g. they are used in arts and real-estate as well as for fishery, spectrum, renewable energy, (online) advertisement, and many more.

Auctions are especially useful when an information asymmetry between buyers and sellers makes setting a fixed price difficult. Ideally, competition between rational bidders leads to an efficient allocation. However, depending on the bidders' value models and the auction mechanism applied, bidders face difficult decision problems about the bid amount to submit. Bayes-Nash equilibria are a central solution concept for these decision problems. However, they are known only for very few and simple auction designs.

In this thesis' first two studies, we considered two large markets, namely the online display ad market and the renewable energy market. We have shown that truthful bidding often is a good strategy in the online display ad market due its large size in terms of bidders and number of items to sell. Assuming truthful bidding for another large market, i.e. the expansion of renewable energy sources, we developed and presented a combinatorial auction design as a viable alternative to the currently applied auction design of selling individual items. The proposed design allows an efficient allocation with almost no increase in remuneration prices and is approximately incentive-compatible in the large. In a third study, we focused on the strategic bidding problem of bidders in general auction designs. We developed an algorithm called Neural Pseudogradient Ascent to learn bidding strategies in a variety of auction games, often leading to Bayes-Nash equilibria. Throughout the numerical experiments, the algorithm has presented itself as a powerful tool for the computation of these equilibria.

With this thesis, we contribute to the literature on bidding strategies in display ad auctions (Sutterer et al., 2019) and to bidding strategies in general auction markets

(Heidekrüger et al., 2019, 2020a,b) as well as to literature on the design of renewable energy auctions (Bichler et al., 2020). The publications herein are not only valuable to the research community but also provide important insights and practical contributions for practitioners in the field.

References

Anatolitis, V., Welisch, M., 2017. Putting renewable energy auctions into action – An agent-based model of onshore wind power auctions in Germany. Energy Policy 110, 394–402. doi:10.1016/j.enpol.2017.08.024.

Arrow, K.J., Debreu, G., 1954. Existence of an equilibrium for a competitive economy. Econometrica: Journal of the Econometric Society , 265–290.

Ausubel, L.M., Cramton, P., Milgrom, P., 2006. The clock-proxy auction: A practical combinatorial auction design. Handbook of Spectrum Auction Design , 120–140.

Azevedo, E.M., Budish, E., 2019. Strategy-proofness in the Large. The Review of Economic Studies 86, 81–116. doi:10.1093/restud/rdy042.

Berg, J., Greenwald, A., Naroditskiy, V., Sodomka, E., 2010. A knapsack-based approach to bidding in ad auctions., in: ECAI, pp. 1013–1014.

Bichler, M., 2017. Market Design: A Linear Programming Approach to Auctions and Matching. Cambridge University Press.

Bichler, M., Ferrell, D., Fux, V., Goeree, J.K., 2019. Designing environmental markets for trading catch shares. INFORMS Journal on Applied Analytics 49, 324–337. URL: https://doi.org/10.1287/inte.2019.1003, doi:10.1287/inte.2019.1003, arXiv:https://doi.org/10.1287/inte.2019.1003.

Bichler, M., Goeree, J.K., 2017. Handbook of Spectrum Auction Design. Cambridge University Press.

Bichler, M., Grimm, V., Kretschmer, S., Sutterer, P., 2020. Market design for renewable energy auctions: An analysis of alternative auction formats. Energy Economics 92, 104904. URL: http://www.sciencedirect.com/science/article/pii/S0140988320302449, doi:https://doi.org/10.1016/j.eneco.2020.104904.

Bosshard, V., Bünz, B., Lubin, B., Seuken, S., 2017. Computing Bayes-Nash Equilibria in Combinatorial Auctions with Continuous Value and Action Spaces, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization, Melbourne, Australia. pp. 119–127. URL: https://www.ijcai.org/proceedings/2017/18, doi:10.24963/ijcai.2017/18.

Bosshard, V., Bünz, B., Lubin, B., Seuken, S., 2020. Computing Bayes-Nash Equilibria in Combinatorial Auctions with Verification. Journal of Artificial Intelligence Research 69, 531–570. URL: http://arxiv.org/abs/1812.01955, doi:https://doi.org/10.1613/jair.1.11525, arXiv:1812.01955.

Brown, G.W., 1951. Iterative solution of games by fictitious play. Activity analysis of production and allocation 13, 374–376.

Chakrabarty, D., Zhou, Y., Lukose, R., 2008. Online knapsack problems, in: Workshop on internet and network economics (WINE).

Chen, J., Stallaert, J., 2014. An economic analysis of online advertising using behavioral targeting. MIS Quarterly 38, 429–449.

Cournot, A.A., 1838. Recherches sur les principes mathématiques de la théorie des richesses. URL: https://gallica.bnf.fr/ark:/12148/bpt6k6117257c.

Daskalakis, C., Goldberg, P., Papadimitriou, C., 2009. The Complexity of Computing a Nash Equilibrium. SIAM Journal on Computing 39, 195–259. URL: https://epubs.siam.org/doi/abs/10.1137/070699652, doi:10.1137/070699652.

Day, R.W., Cramton, P., 2012. Quadratic Core-Selecting Payment Rules for Combinatorial Auctions. Operations Research 60, 588–603. URL: https://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1024, doi:10.1287/opre.1110.1024.

Day, R.W., Raghavan, S., 2007. Fair Payments for Efficient Allocations in Public Sector Combinatorial Auctions. Management Science 53, 1389–1406. URL: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1060.0662, doi:10.1287/mnsc.1060.0662.

De, L., Chen, J., Whinston, A.B., 2010. Ex ante information and the design of keyword auctions. Information Systems Research 21, 133–153.

Dixon, H.D., 2001. Equilibrium and explanation, in: Surfing Economics. Springer, pp. 21–53.

emarketer, 2019. Global digital ad spending 2019. https://www.emarketer.com/content/global-digital-ad-spending-2019 Accessed on 2020-09-21.

Fudenberg, D., Levine, D.K., 1999. The Theory of Learning in Games. volume 2 of *MIT Press Series on Economic Learning and Social Evolution*. 2. ed., MIT Press, Cambridge.

Fudenberg, D., Levine, D.K., 2009. Learning and Equilibrium. Annual Review of Economics URL: https://dash.harvard.edu/handle/1/4382413, doi:10.1146/annurev.economics.050708.142930.

Gallachóir, B.P.Ó., Bazilian, M., McKeogh, E.J., 2009. Wind-Energy Policy Development in Ireland: A Critical Analysis, in: Stachan, P., Lal, D., Toke, D. (Eds.), Wind Power and Power Politics. Routledge, New York, pp. 128–155.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press. `http://www.deeplearningbook.org`.

Grashof, K., 2013. Herausforderungen bei Ausschreibungsverfahren für Onshore-Windenergie und PV, in: Leprich, U., Grashof, K., Guss, H., Klann, U., Weber, A., Zipp, A. (Eds.), Stromsystem-Design: das EEG 2.0 und Eckpfeiler eines zukünftigen Regenerativwirtschaftsgesetzes. IZES, Saarbrücken.

Grimm, V., Sölch, C., Zöttl, G., 2017. Regionalkomponenten bei der EE-Vergütung. Technical Report. Report on behalf of the German Monopoly Commission. URL: `http://www.wirtschaftstheorie.wiso.uni-erlangen.de/de/forschung/gutachten/`.

Hauser, E., Kochems, J., 2014. Ausschreibungsmodelle für Wind Onshore: Erfahrungen im Ausland. Technical Report. IZES. Saarbrücken.

Hauser, E., Weber, A., Zipp, A., Leprich, U., 2014. Bewertung von Ausschreibungsverfahren als Finanzierungsmodell für Anlagen erneuerbarer Energienutzung. Technical Report. IZES. Saarbrücken.

Heidekrüger, S., Sutterer, P., Bichler, M., 2019. Computing approximate bayes-nash equilibria through neural self-play, in: Workshop on Information Technology and Systems (WITS19), Munich, Germany.

Heidekrüger, S., Sutterer, P., Kohring, N., Bichler, M., 2020a. Equilibrium learning in combinatorial auctions: Computing approximate bayesian nash equilibria via pseudo-gradient dynamics, in: Workshop on Information Technology and Systems (WITS20), Hyderabad, India.

Heidekrüger, S., Sutterer, P., Kohring, N., Bichler, M., 2020b. Learning bayesian nash equilibria in auction games, in: INFORMS Workshop on Data Science, Maryland, USA.

Kingma, D.P., Ba, J., 2017. Adam: A method for stochastic optimization. arXiv:1412.6980.

Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S., 2017. Self-normalizing neural networks, in: Advances in neural information processing systems, pp. 971–980.

Krishna, V., 2009. Auction Theory. Academic press.

McLean, A., McGovern, P., Donelly, K., 2007. Ireland. International Energy Law and Taxation Review 10, 207–210.

Nash, J.F., et al., 1950. Equilibrium points in n-person games. Proceedings of the national academy of sciences 36, 48–49.

Nisan, N. (Ed.), 2007. Algorithmic Game Theory. Cambridge University Press, Cambridge ; New York. OCLC: ocn122526907.

Perlich, C., Dalessandro, B., Hook, R., Stitelman, O., Raeder, T., Provost, F., 2012. Bid optimizing and inventory scoring in targeted online advertising, in: Yang, Q., Agarwal, D., Pei, J. (Eds.), Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '12, ACM Press, New York, New York, USA. p. 804. doi:10.1145/2339530.2339655.

Qian, N., 1999. On the momentum term in gradient descent learning algorithms. Neural networks 12, 145–151.

del Río, P., 2017. Designing auctions for renewable electricity support. Best practices from around the world. Energy for Sustainable Development 41, 1–13. doi:10.1016/J.ESD.2017.05.006.

del Río, P., Linares, P., 2014. Back to the future? Rethinking auctions for renewable electricity support. Renewable and Sustainable Energy Reviews 35, 42–56. doi:10.1016/j.rser.2014.03.039.

Salimans, T., Ho, J., Chen, X., Sidor, S., Sutskever, I., 2017. Evolution Strategies as a Scalable Alternative to Reinforcement Learning. arXiv:1703.03864 [cs, stat] URL: http://arxiv.org/abs/1703.03864, arXiv:1703.03864.

Shapley, L., 1964. Some topics in two-person games. Advances in game theory 52, 1–29.

Sutterer, P., Waldherr, S., Bichler, M., 2019. Are truthful bidders paying too much? efficiency and revenue in display ad auctions. ACM Transactions on Management Information Systems 10, 1–18. doi:10.1145/3325523.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. A Bradford Book, Cambridge, MA, USA.

Tiedemann, S., Förster, S., Wigand, F., 2016. Auctions for Renewable Energy Support in Italy: Instruments and lessons learnt. Technical Report. AURES Report D4.1 - IT. URL: http://auresproject.eu/sites/aures.eu/files/media/countryreports/pdf{_}italy.pdf.

Vickrey, W., 1961. Counterspeculation, auctions, and competitive sealed tenders. The Journal of finance 16, 8–37.

Wigand, F., Förster, S., Amazo, A., Tiedemann, S., 2016. Auctions for Renewable Energy Support: Lessons Learnt from International Experiences. Technical Report. AURES Report D4.2. URL: http://auresproject.eu/publications/auctions-renewable-energy-support-lessons-learnt-from-international-experiences.

Winkler, J., Magosch, M., Ragwitz, M., 2018. Effectiveness and efficiency of auctions for supporting renewable electricity – what can we learn from recent experiences? Renewable Energy 119, 473 – 489. URL: http://www.sciencedirect.com/science/article/pii/S0960148117309357, doi:https://doi.org/10.1016/j.renene.2017.09.071.

Yang, S., Ghose, A., 2010. Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? Marketing Science 29, 602–623.

Zhang, W., Wang, J., 2015. Statistical arbitrage mining for display advertising, in: Cao, L., Zhang, C., Joachims, T., Webb, G., Margineantu, D.D., Williams, G. (Eds.), Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15, ACM Press, New York, New York, USA. pp. 1465–1474. doi:10.1145/2783258.2783269.

Zhang, W., Yuan, S., Wang, J., 2014. Optimal real-time bidding for display advertising, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA. p. 1077–1086. URL: https://doi.org/10.1145/2623330.2623633, doi:10.1145/2623330.2623633.

Zhang, W., Yuan, S., Wang, J., Shen, X., 2015. Real-time bidding benchmarking with ipinyou dataset. arXiv:1407.7073.

Zhao, X., Xue, L., 2012. Competitive target advertising and consumer data sharing. Journal of Management Information Systems 29, 189–222.