NORTHWESTERN UNIVERSITY

Approximate Optimality of Simple Mechanisms

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Yingkai Li

EVANSTON, ILLINOIS

June 2022

ABSTRACT

Approximate Optimality of Simple Mechanisms

Yingkai Li

We consider general utility models and information structures of the agents and illustrate when economic conclusions for designing simple mechanisms in classical settings extends for general environments. We show that whether economic conclusions can be generalized depends on the details of the generalizations. For example, in single-item auction, competition and non-anonymity are not crucial factors for revenue maximization when agents have linear utilities [Yan, 2011, Alaei et al., 2018], and these conclusions extend for broad classes of non-linear utilities. In comparison, the economic conclusions we derived for exogenous information settings often fail when the information is endogenous. For example, in multi-dimensional information acquisition problems, scoring the agent separately is without loss when the signals are exogenous, but suffers a great loss when the signals are endogenous. In selling information problems, price discrimination and commitment to revealing partial information are crucial for revenue maximization if the agent has an exogenous signal about the unknown state [Bergemann et al., 2021]. However, pricing for full information is approximately optimal when the signal is endogenous.

Acknowledgements

I want to thank my advisor Jason Hartline for his guidance during my PhD study. It has been a great pleasure to work with him for these years, and I have learned a lot from him, including but not limited to how to develop a good taste for researches, how to solve hard math questions, or how to communicate with other researchers to convey the novelty of our results. These great experience will continue to be helpful for me when I pursue my career in academics in the future.

I also want to express my gratitude to Harry Pei, who essentially acted as my second advisor from the economic department. Without his help, I wouldn't be able to discover many interesting topics that economists are excited about, or learn how to think as an economist when conducting economic researches. He really helped me enter the greater domain of economics, which eventually lead me to accomplish great work beyond the classical computer science perspective.

I am also grateful to Brendan Lucier, Nicole Immorlica, Alex Slivkins and Moshe Babaioff for their hospitality when I visited Microsoft Research for internship during the summer of 2020 and 2021. This is the first time I had an internship at an industrial research lab, and it is really exciting for me to see how theory is conducted in a more applied perspective and connected to real-world applications.

Moreover, I want to thank all my dissertation committee members, Jason Hartline, Harry Pei, Brendan Lucier, Yannai Gonczarowski and Annie Liang. My thesis has benefited a lot from the discussions with them, and their valuable comments and suggestions.

I'm also very fortunate to have worked with many other brilliant researchers in this field: Bo Li, Yaonan Jin, Pinyan Lu, Yuan Zhou, Yiding Feng, Jing Chen, Xiaotie Deng, Edmund Lou, Liren Shan, Yifan Wu, Aleck Johnsen, Yining Wang, Jason Gaitonde, Bar Light, Haoran Ye, Kefan Dong, Qin Zhang, Xiaowei Wu, Inbal Talgam-Cohen, Tal Alon, Paul Dütting, Hedyeh Beyhaghi, Matt Weinberg, Haifeng Xu, and Boli Xu. I would not be able to write many nice papers without your help.

It has been a wonderful time for me, both academically and personally, to spend my PhD at Northwestern University. It is pleasure to be a part of this amazing community, which includes many great faculties: Samir Khuller, Ben Golub, Konstantin Makarychev, Aravindan Vijayaraghavan and Yingni Guo; and many of my colleges and friends: Modibo Camara, Aidao Chen, Xue Chen, Saba Ahmadi, Huck Bennett, Abhratanu Dutta, Sanchit Kalhan, Jinshuo Dong, Matthew vonAllmen, Shirley Zhang, Michalis Mamakos, Sheng Long, Aravind Reddy, Leif Rasmussen, Sheng Yang, Chenhao Zhang, Charles Cui, Anant Shah, Jingyuan Wang, Anran Li and Xiaoyun Qiu.

Finally and mostly importantly, I would like to thank my parents, Tianwang Li and Lianying Wang, my wife Wenyang Li, and all other family members. I could not accomplish any of this without your love and support.

Table of Contents

List of Figures

CHAPTER 1

Introduction

The *method of approximation* quantifies the extent to which a theory can be generalized from ideal models and enables the separation of details from salient features of the model. Given an (possibly complicated, detail-dependent) optimal mechanism for an objective like revenue, there may exist other (simple, detail-free) mechanisms that *approximate* (i.e., attain a "large" fraction of) the optimal revenues. In this case we may say that the theory behind the simple mechanisms generalizes from the ideal model. Otherwise, it does not.[1]

There are extensive studies of simple mechanisms with approximation guarantees for the classical mechanism design problems with specific assumptions on the agents' utility models. For example, Yan [2011] shows that when agents have linear utilities, sequential posted pricings, which arrange the agents in an order and offer while-supplies-last posted prices, guarantee an $e/(e-1)$-approximation, i.e., the best order and prices achieves at least 63.2% of the optimal auction revenue. This implies that simultaneity and competition are not necessary drivers for revenue maximization for linear utility agents. Another example is information acquisition, in order to elicit high-dimensional information, it is often without loss to elicit marginal information on each dimensional separately [Lambert, 2011]. In the model of selling information, Bergemann et al. [2021] show that when the information of the agents is exogenous, pricing for full information cannot guarantee any

[1]See the survey of Hartline [2012] for detailed discussion of the method of approximation in economics.

non-trivial fraction of the optimal revenue. This implies that price discrimination is a crucial factor for revenue maximization given exogenous information.

We study the problem that to what extend the economic conclusions from the classical mechanism design literature generalize for agents with general utility models and general information structures. In particular, in this thesis, we focus on two specific generalizations: from linear utilities to non-linear utilities, and from exogenous information to endogenous information. We show that many economic lessons we obtained for linear utilities generalize for broad class of non-linear utilities, while the economic lessons we obtained for exogenous information fail for endogenous information. Therefore, there is no unified solutions for generalizing the economic conclusions to complex models, and more investigation is required for understanding the performance of simple mechanisms in various general environments. In the next section, we will summarize the main results in this thesis.

1.1. Main Contributions

1.1.1. Non-linear Utilities

For classical auction design for agents with linear utilities, Bulow and Roberts [1989] show that the marginal revenue maximization mechanism is revenue optimal, drawing a close connection between the classical microeconomics and the auction theory. Yan [2011] shows that sequential posted pricings guarantee an $e/(e-1)$-approximation, which suggests the relative irrelevance of simultaneity and competition for revenue maximization. Jin et al. [2019] further show that when the agents' (non-identical) value distributions satisfy a concavity property, a.k.a., "regular distributions", posting an anonymous price guarantees a 2.62-approximation, which implies that discrimination across different agents is not essential either.

We generalize these approximation results from linear agents to non-linear agents.[2] From this generalization, not only do we observe that the main drivers of good mechanisms are similar for non-linear agents, but also that non-linearity itself is not a main concern that necessitates specialized mechanism designs (beyond the approach of our generalization).

Bulow and Roberts [1989], as later interpreted by Alaei et al. [2013], show that to design optimal mechanisms for linear agents, it is without loss to restrict attention to *pricing-based mechanisms*, i.e., mechanisms where the menu offered to each agent is equivalent to a distribution over posted prices. The multi-agent mechanism design problem can be decomposed as single-agent mechanism design problems through the reduced-form approach of Border [1991]. From Bulow and Roberts [1989], the solution to these single-agent problems for linear agents are (possibly randomized) price postings and the optimal mechanism can be interpreted as marginal revenue maximization. Thus, every mechanism for linear agents is equivalent to a pricing-based mechanism.

Pricing-based mechanisms can be generalized to non-linear agents by considering *per-unit* prices, i.e., given per-unit price p, an agent can purchase any lottery with winning probability $q \in [0, 1]$ and pay price $p \cdot q$ in expectation. For non-linear agents (e.g., agents with budget constraints), not all mechanisms can be interpreted as pricing-based mechanisms and, in fact, pricing-based mechanisms are not generally optimal. Nonetheless, we show that these mechanisms are approximately optimal for large families of non-linear

[2] In this thesis, we write "agents with linear utilities" as "linear agents" for short, and "agents with non-linear utilities" as "non-linear agents".

agents. For these families we say that the non-linear agents *resemble* linear agents. We introduce a reduction framework. Given a pricing-based mechanism that guarantees a β-approximation (i.e., achieves at least $1/\beta$ fraction of the optimal objective) for linear agents and given non-linear agents that are ζ-*resemblant*[3] of linear agents and satisfy the von Neumann-Morgenstern expected utility representation [Morgenstern and von Neumann, 1953], the reduction framework transforms the aforementioned pricing-based mechanism for linear agents into an analogous pricing-based mechanism for the non-linear agents. The non-linear agent mechanism guarantees a $\beta\zeta$-approximation bound.

The reduction framework can be combined with approximation results for linear agents to show that simple mechanisms such as marginal revenue maximization, sequential posted pricing, and anonymous pricing are approximately optimal for non-linear agents that resemble linear agents, and the economic lessons (e.g., non-cruciality of simultaneity, competition, discrimination) derived from those mechanisms for linear agents can be lifted to non-linear agents. As an example, agents with independent private budget and regular valuation distribution are 3-resemblant of linear agents, which implies that the approximation of sequential posted pricing for such non-linear agents is $3e/(e-1)$.

This thesis characterizes broad families of non-linear agents that are ζ-resemblant for small constant factors ζ (e.g., agents with independent private budget and regular valuation distribution) and families that are not (e.g., agents whose budget and value are correlated). For non-linear agents that are ζ-resemblant, pricing-based mechanisms are approximately optimal wherever they are approximately optimal for linear agents; thus, non-linearity of utility can be viewed as a detail that can be omitted from the

[3]We measure the resemblance of agents in terms of the (topological) closeness of their revenue curves, as defined in Bulow and Roberts [1989]. We provide the details in Chapter 2.

model without significantly altering the main take-aways. On the other hand, with utility models that are not ζ-resemblant for modest ζ, non-linearity is a crucial feature that needs specific study for identifying forms of mechanisms lead to good economic outcomes.

Our reduction framework can be applied more broadly for non-linear agents beyond the expected utility theory with the restriction to *posted pricing mechanisms*[4] (e.g., sequential posted pricing, anonymous pricing). For instance, when agents have stochastic outside options – which can be viewed as a special form of non-linear utility that does not satisfy expected utility theory – are 2-resemblant under a concavity assumption. Thus, for such agents, sequential posted pricing is approximately optimal and the economic lessons from previous discussions generalize.

1.1.2. Information Acquisition

We consider the problem of an uniformed principal acquiring information from a strategic agent. In particular, we formalize the problem as optimizing scoring rules for reporting the expectation of a incentivizing the agent to exert effort on acquiring additional information.

Proper scoring rules incentivize a forecaster to reveal her true belief about an unknown and probabilistic state. The principal publishes a scoring rule that maps the reported belief and the realized state to a reward for the forecaster. The forecaster reports her belief about the state. The state is realized and the principal rewards the forecaster according to the scoring rule. A scoring rule is proper if the forecaster's optimal strategy, under any belief she may possess, is to report that belief. Proper scoring rules are also designed for directly eliciting a statistic of the distribution such as its expectation.

[4]Posted pricing mechanisms are pricing-based mechanisms where prices posted to each agent do not depend on actions of other agents.

Not all proper scoring rules work well in any a given scenario. This thesis considers a mathematical program for optimization of scoring rules where (a) the objective captures the incentive for the forecaster to exert effort and (b) the boundedness constraints prevent the principal from scaling the scores arbitrarily. For (a), we focus on a simple binary model of effort where the forecaster does or does not exert effort and with this effort the forecaster obtains a refined posterior distribution from the prior distribution on the unknown state (e.g., by obtaining a signal that is correlated with the state). We adopt the objective that takes the perspective of the forecaster at the point of the decision with knowledge of both the prior and the distributions of posteriors that is obtained by exerting effort. We want a scoring rule that maximizes the difference in expected scores for the posterior distribution and prior distribution. For (b), we impose the ex post constraint that the score is in a bounded range, i.e., without loss, between zero and one. Notice that this program would be meaningless without a constraint on the scores - otherwise the score could be scaled arbitrarily - and it would be meaningless without considering the difference in scores between posterior and prior - otherwise any bounded scoring rule scaled towards zero plus a constant close to the upper bound would be near optimal.

We solve for the optimal scoring rule for reporting the expectation in single-dimensional space. As we expect for single-dimensional mechanism design problems for an agent with linear utility [Myerson, 1981b], the optimal scoring rule is a step function (which induces a V-shaped scoring rule with its lower tip at the expectation of the prior belief). To implement this V-shaped scoring rule, it is sufficient for the designer to know the prior mean instead of the details on the distribution over posteriors. We also demonstrate a first result for prior-independent analysis of scoring rules. Among scoring rules for reporting the expectation, the quadratic scoring rule is within a constant factor of optimal.

For multi-dimensional forecasting, without concern of acquiring additional information, a simple choice of the principal is to elicit information separately across different dimensions, and the aggregated information would still be accurate.

When the agent can acquire a costly signal, we show that the gap between acquiring information separately and optimally can be linear in the size of the dimensions. For symmetric distributions, we give an analytical characterization of the optimal scoring rule as inducing a V-shaped utility function. For multi-dimensional forecasting without a symmetry assumption, we identify a V-shaped scoring rule that gives an 8-approximation. This scoring rule can be interpreted as scoring the dimension for which the agent's posterior in the optimal single-dimensional scoring rule gives the highest utility. Equivalently, it can be implemented by letting the agent select which dimension to score and only scoring that dimension (after exerting effort to learn the posterior mean of all dimensions). Moreover, while optimal mechanisms generally depend on the distribution over posteriors, our approximation bounds are proved for simple mechanisms (V-shaped scoring rules) that depend only on the prior mean, and do not require detailed knowledge of the distribution over posteriors.

1.1.3. Selling Information

We consider the problem of maximizing the revenue of the data broker, where the agent can endogenously acquire additional information. Specifically, there is an unknown state and both the data broker and the agent have a common prior over the set of possible states.

The data broker can offer a menu of information structures for revealing the states with associated prices to the agent. Then the agent picks the expected utility maximization entry from the menu, and pays the corresponding price to the data broker. The agent has a private valuation for information and can acquire additional costly information upon receiving the signal from the data broker. The literature has acknowledged the possibility for the agents to conduct their own experiments to be privately informed of the states [e.g., Bergemann, Bonatti, and Smolin, 2018]. The distinct feature in our model is that the decision for acquiring additional information is endogenous. Specifically, after receiving the signal from the data broker, the agent can subsequently acquire additional information with costs. For example, the agent is a decision maker who chooses an action to maximize her expected utility based on her posterior belief over the states. The agent will first acquire information from the data broker, and based on her posterior, she can potentially conduct more experiments to refine her belief before taking the action. Another example captured in our model is where the agent is a firm that sells products to consumers, and the information the data broker provides is a market segmentation of the consumers. The firm has a private and convex cost for producing different level of qualities for the product, and the firm can conduct his own experiments (e.g., sending surveys to potential consumers) with additional costs to further segment the market after receiving the information from the data broker.[5] In addition, in our model, we allow the firm to repeat the market research until it is not beneficial to do so, i.e., when the cost of information exceeds the marginal benefits of information. This captures the situation

[5] Yang [2020] studies a similar model, where the firm cannot conduct his own market research to refine his knowledge. Moreover, the cost function of the firm is linear in Yang [2020], which leads to a qualitatively different result compared to our model. See Section 4.3 for a detailed discussion.

that the firm can decide the date for announcing the product to the market, and before the announcement, the firm sends out surveys to potential consumers each day to learn the segmentation of the markets. At the end of each day, the firm receives an informative signal through the survey, and decides whether to continue the survey in next day, or stop the survey and announce the product with corresponding market prices to the public.

When the private information of the agent is exogenous, [e.g., Bergemann, Bonatti, and Smolin, 2018] show that in the revenue optimal mechanism, the menu complexity can be linear in the number of action choices of the agent in the worst case, and posting a deterministic price for revealing full information cannot guarantee a constant approximation to the optimal. This suggests that third degree price discrimination is crucial for revenue maximization in exogenous information setting.

In this model of allowing the agent to acquire information endogenously, we impose a linearity assumption on the agent's private preference over different experiments, i.e., the value of the agent for any posterior distribution is simply the product of her private type and the value of the posterior distribution. In the examples we provided in previous paragraphs, both the decision maker who chooses an optimal action to maximize her payoff based on the posterior belief and the firm that sells products to consumers to maximize the revenue satisfy the linear valuation assumption. Essentially, this condition assumes that the private type of the agent represents her value for additional information, and there is a linear structure on the preference. It excludes the situation where the private type of the agent represents an exogenous private signal correlated with the states. We show that with linear valuations, when the agent can acquire additional costly information, there exists a threshold type θ^* such that (1) for any type $\theta \geq \theta^*$, the optimal mechanism reveals full information to the agent; and (2) for any type $\theta < \theta^*$, the optimal mechanism may reveal partial information and the individual rational constraint always binds. The first statement is the standard no distortion at the top observation in the optimal mechanisms. The second statement suggests that the optimal mechanism may discriminate lower types of the agent by offering the options of revealing partial information to the agent with lower prices. Moreover, the allocations and the prices for those lower types are set such that the agent is exactly indifferent between participation and choosing the outside option (by conducting her own experiments with additional costs). Our characterizations suggest that the optimal mechanisms for selling information may be complex and contain a continuum of menu entries when the information is endogenous. However, posted pricing for revealing full information achieves at least half of the optimal revenue in the worst case. This suggests that price discrimination is not crucial for approximating the optimal revenue in endogenous information setting, which leads to a sharp contrast to the exogenous information setting.

1.2. Related Work

Non-linear Utilities. Frameworks for reducing approximation for non-linear agents to approximation for linear agents has also been studied in Alaei et al. [2013]. This reduction framework converts the marginal revenue mechanism for linear agents to a mechanisms for non-linear agents and general objectives. Their reduction framework is also applicable to other DSIC, IIR, deterministic mechanisms for linear agents. Unlike our framework which uses single-agent price-posting mechanisms (induced from price-posting payoff curves) as a building-block, Alaei et al. [2013] convert mechanisms for linear agents into mechanisms for non-linear agents with single-agent ex ante optimal mechanisms (induced from optimal payoff curves) as components. From the mechanism designer's perspective, identifying ex ante optimal mechanisms for a single non-linear agents can be much harder than identifying ex ante optimal price-posting mechanisms (e.g., private budget utility, risk averse utility). Furthermore, due to this difference, the implementation of the reduction framework together with its outcome mechanisms in Alaei et al. [2013] is more complex than ours. In general, the framework in Alaei et al. [2013] converts DSIC mechanisms for linear agents into Bayesian incentive compatible mechanisms for non-linear agents.

Mechanism design for non-linear agents is well studied in the literature. In this work, as applications of our general framework, we focus on three specific non-linear models, agents with budget constraints, agents with risk averse attitudes, and agents with endogenous valuation.

Laffont and Robert [1996] and Maskin [2000] study the revenue-maximization and welfare-maximization problems for symmetric agents with *public* budgets in single-item environments. Boulatov and Severinov [2018] generalize their results to agents with i.i.d. values but asymmetric public budgets. Che and Gale [2000] consider the single agent problem with *private* budget and valuation distribution that satisfies declining marginal revenues, and characterize the optimal mechanism by a differential equation. Devanur and Weinberg [2017a] consider the single agent problem with private budget and an arbitrary valuation distribution, characterize the optimal mechanism by a linear program, and use an algorithmic approach to construct the solution. Pai and Vohra [2014] generalize the characterization of the optimal mechanism to symmetric agents with uniformly distributed private budgets. Richter [2019] shows that a price-posting mechanism is optimal for selling a divisible good to a continuum of agents with private budgets if their valuations are regular with decreasing density. For more general settings, no closed-form characterizations are known. However, the optimal mechanism can be solved by a polynomial-time solvable linear program over interim allocation rules [cf. Alaei et al., 2012, Che et al., 2013].

Most results for agents with risk-averse utilities consider the comparative performance of the first- and second-price auctions, cf., Holt Jr [1980], Che and Gale [2006]. Matthews [1983] and Maskin and Riley [1984], however, characterize the optimal mechanisms for symmetric agents for constant absolute risk aversion and more general risk-averse models. Baisa [2017] shows that the optimal mechanism for risk averse agents departs from the linear agents, since the optimal mechanism does not allocate to the highest bidder, and can better screen the agents through allocating the item to a group of agents with lotteries. Gershkov et al. [2021b] show that if the seller can make positive transfer to the agents, the optimal mechanism features the property that under equilibrium, all agents face no uncertainty in the realized utility.

The model for agents with endogenous valuation has been studied extensively in Tan [1992], King et al. [1992], Gershkov et al. [2021a], Akbarpour et al. [2021] where agents can make costly investment before the auction. This is a generalization of the model for agents with entry costs [Celik and Yilankaya, 2009]. This main focus of the literature is to characterize the optimal mechanisms in restricted settings. For example, Gershkov et al. [2021a] characterize the revenue optimal symmetric mechanism for symmetric buyers.[6] The reduction framework in this thesis implies that sequentially offering a price to each

[6] Gershkov et al. [2021a] also showed that even for symmetric buyers, symmetric mechanism may not be revenue optimal among all possible mechanisms.

agent is a constant approximation for both welfare and revenue maximization when there are multiple asymmetric buyers. Akbarpour et al. [2021] consider approximating the optimal welfare when it is computationally intractable to find the optimal allocation. They show that any algorithm that excludes bossy negative externalities can be converted to a mechanism that guarantees the same approximation ratio to the optimal welfare. They restrict attention to full information equilibrium, while our analysis applies to settings with private valuations.

It is well known that simple mechanisms generate robust performance guarantees for both welfare maximization [Roughgarden et al., 2017] and revenue maximization [Carroll, 2017, Bei et al., 2019]. Moreover, simple mechanisms are approximately optimal under natural assumptions of type distributions. For single item auction and linear agents, Jin et al. [2019] show that the tight ratio between anonymous pricing and the optimal mechanism is 2.62 under regularity assumption, and Yan [2011] shows that the tight approximation ratio is ${}^{e}/_{(e-1)}$ for sequential posted pricing. The approximate optimality of sequential posted pricing can be generalized to multi-item settings when agents have unit-demand valuations [Chawla et al., 2010, Cai et al., 2016]. For non-linear agents, given matroid environments, Chawla et al. [2011] show that a simple lottery mechanism is a constant approximation to the optimal pointwise individually rational mechanism for agents with monotone-hazard-rate valuations and private budgets. In contrast, our approximation results are with respect to the optimal mechanism under interim individually rationality which can be arbitrarily larger than the benchmark from Chawla et al. [2011]. For multiple items, Cheng et al. [2018] shows that selling items separately or as a bundle is approximately optimal for a single agent with additive valuation. Our analyses uses one of their lemmas.

Scoring Rules. Characterizations of scoring rules for eliciting the mean and for eliciting a finite-state distribution play a prominent role in our analysis. Previous works show, in various contexts, that scoring rules are proper if and only if their induced utility functions are convex. McCarthy [1956] characterized proper scoring rules for eliciting the full distribution on a finite set of states. Osband and Reichelstein [1985] characterized continuously differentiable scoring rules that elicit multiple statistics of a probability distribution. Lambert [2011] characterized the statistics that admit proper scoring rules and characterized the uniformly-Lipschitz-continuous scoring rules for the mean of a single-dimensional state. Abernethy and Frongillo [2012] characterized the proper scoring rules for the marginal means of multi-dimensional random states in the interior of the report space. We augment this characterization by showing that the induced utility function converges to a limit on the boundary of the report space. This augmentation enables us to write the mathematical program that optimizes over the whole report space.

Most of the prior work looking at incentives of eliciting information considers a fundamentally different model from ours. This prior work typically focuses on the incentives of the forecaster to exert effort to obtain a signal (a.k.a., a data point), but then assumes that this data point is reported directly (and cannot itself be misreported). In this space, Cai, Daskalakis, and Papadimitriou [2015] considers the learning problem where the principal aims to acquire data to train a classifier to minimize squared error less the cost of eliciting the data points from individual agents. The mechanism for soliciting the data from the agents trades off cost (in incentivizing effort) for accuracy of each individual point. Chen,