Setting	VCG		RegretNet (truthful)		RegretNet (adversarial)	
	utility	*revenue*	*utility*	*revenue*	*utility*	*revenue*
Setting Ia : two bidders, two objects uniform value distribution	0.336	0.666	0.149	0.882	0.306 (+108%)	0.696 (-21%)
Setting Ib : two bidders, two objects exponential value distribution	1.000	1.000	0.504	1.481	0.574 (+13%)	1.443 (-2.5%)
Setting IIa : three bidders, two objects uniform value distribution	0.166	1.000	0.096	1.034	0.148 (+54%)	0.985 (-4.7%)
Setting IIb : three bidders, two objects exponential value distribution	0.666	1.666	0.249	1.804	0.294 (+18%)	1.801 (-0.1%)

Table 2: Experiments for the multi-item setting. The strategic bidder is using a linear bidding exploration policy with parameter $\sigma_k^2 = 0.05$. The seller is using the RegretNet architecture as selling mechanism. We run $T = 150$ adversarial training epochs,and base our evaluation on averaging over $q = 12$ strategies from the exploration policy.

5.1 Experiments with multi-item auctions

Using simple linear shadings in multi-item settings yielded considerable improvements in bidders' utility. We implemented Algorithm 1 initializing μ_1 to be an array of m ones (corresponding to the thuthful strategy), and $\sigma_k^2 = 0.05$ for all $k \in M$. We run $T = 150$ adversarial training epochs, and sample $q = 12$ lambdas per epoch. We optimize the seller mechanism every 3 adversarial epoch by training the RegretNet architecture. We implement the RegretNet architecture in PyTorch by using two neural networks with $H = 2$ hidden layers of size $h = 30$. Our experimental results are reported in Table 2. We observe substantial improvements in bidders' utility, with a 108% uplift for Setting Ia and a 54% uplift for Setting IIa. This is the performance of the exploration and it would be possible to improve the strategic bidder's utility by decreasing the variance of the exploration policy at the cost of not being robust to changes of the learning mechanism. This suggests that even better improvements in utility could be found using more complex bidding strategies in the spirit of the thresholded-virtual-value strategy introduced by [28] for the single-item framework.

Our work thus opens the door to several natural extensions such as using neural networks to parametrize more complex bidding strategies, or studying other bidder types, valuation distributions and auctions such as the combinatorial auction. However, training neural networks to learn the exploration policy would increase the running time of the procedure, which is already substantial for linear shading strategies. This provides a first benchmark to design adversarial attacks against sellers' learning algorithms. This benchmark could be extended in the near future by testing new seller algorithms and new architecture to learn strategic behaviors. This reinforces the idea that the conceptual mistake of not treating the game where the seller uses past bids to optimize the auction as a Stackelberg game can be very costly for bidders. Moreover, they show that data-driven automatic mechanisms are vulnerable to adversarial attacks, hence providing motivation for practical implementation of adversarial attacks on modern marketplaces, or implementation of automatized mechanisms robust to adversarial attacks on these same platforms.

6 A NEED FOR ADVERSARIALLY-ROBUST SELLER LEARNING MECHANISMS

A natural extension to the design of adversarial attacks against data-driven automated selling mechanisms is the design of learning algorithms which are robust to adversarial attacks. This line of work has been initiated by [2], who find mechanisms which maximize the seller's revenue against the worst bid distribution in a certain class. To avoid dealing with worst-case scenarii, an intermediate approach would be to consider mechanisms robust to a class of bidding strategies and a class of initial value distributions.

Definition 5 (ϵ adversarially-robust learning algorithm). *A selling learning algorithm $\mathcal{M}$ is said to be ϵ adversarially-robust for this class of value distributions, if for any value distributions F_i in this class, for any adversarial attack β^*, with ,β^{Tr} the truthful strategy, the seller's revenue R when the strategic bidder is using U verifies* $R(\mathcal{M}(F_i, \beta^*), \beta^*) \geq R(\mathcal{M}(F_i, \beta^{Tr}),, \beta^{Tr}) - \epsilon$.

This leads to a new definition of incentive compatible learning algorithms where bidders have an incentive to bid truthfully even if the seller is using past bids to optimize her mechanism. A follow up on our work could be to investigate feasibility of such robust mechanisms by adding a constraint to an augmented Lagrangian method similar to that used by [17]. Our approach is the first necessary step in the design of such robust mechanisms since it computes how the revenue is impacted when using a given learning mechanism.

7 CONCLUSION

We present a new way to design adversarial attacks against cutting-edge automatic mechanism design algorithms. Our approach yields very substantial utility gains for the strategic bidder in our numerical experiments. This allows buyers to quantify the price of revealing information about their values in repeated auctions. From a theoretical standpoint, this offers a new tool to study economics interactions through an algorithmic lens and represents a new step to reinterpret economics problems as algorithmic learning problems between strategic agents.

REFERENCES

[1] Michael Albert, Vincent Conitzer, and Peter Stone. 2017. Automated design of robust mechanisms. In *Proceedings of AAAI*.

[2] Amine Allouah and Omar Besbes. 2018. Prior-Independent Optimal Auctions. In *Proceedings of EC*.

[3] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. [n.d.]. Learning prices for repeated auctions with strategic buyers, In 2013. *Proceedings of NIPS*.

[4] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2014. Repeated contextual auctions with strategic buyers. In *Proceedinga of NIPS*.

[5] Mark Armstrong. 1996. Multiproduct nonlinear pricing. *Econometrica* 64, 1 (1996), 51.

[6] Itai Ashlagi, Constantinos Daskalakis, and Nima Haghpanah. 2016. Sequential mechanisms with ex-post participation guarantees. In *Proceedings of EC*.

[7] Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. 2018. A general theory of sample complexity for multi-item profit maximization. In *Proceedings of EC*.

[8] Santiago R Balseiro, Ozan Candogan, and Huseyin Gurkan. 2020. Multistage Intermediation in Display Advertising. *Manufacturing & Service Operations Management* (2020).

[9] Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. 2018. Selling to a no-regret buyer. In *Proceedings of EC*.

[10] Yang Cai, Constantinos Daskalakis, and Christos Papadimitriou. 2015. Optimum statistical estimation with strategic data sources. In *Proceedings of COLT*.

[11] Richard Cole and Tim Roughgarden. 2014. The sample complexity of revenue maximization. In *Proceedings of Theory of computing*.

[12] Vincent Conitzer and Tuomas Sandholm. 2002. Complexity of mechanism design. In *Proceedings of UAI*.

[13] Vincent Conitzer and Tuomas Sandholm. 2006. Computing the optimal strategy to commit to. In *Proceedings of EC*.

[14] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. 2013. Mechanism design via optimal transport. In *Proceedings of the fourteenth ACM conference on Electronic commerce*. ACM, 269–286.

[15] Yuan Deng, Jon Schneider, and Balasubramanian Sivan. 2019. Prior-Free Dynamic Auctions with Low Regret Buyers. In *Proceedings of NeurIPS*.

[16] Mahsa Derakhshan, Negin Golrezaei, and Renato Paes Leme. 2019. LP-based Approximation for Personalized Reserve Prices. *Proceedings of EC* (2019).

[17] Paul Dütting, Zhe Feng, Harikrishna Narasimhan, and David C Parkes. 2019. Optimal auctions through deep learning. In *Proceedings of ICML*.

[18] Alessandro Epasto, Mohammad Mahdian, Vahab Mirrokni, and Song Zuo. 2018. Incentive-aware learning for large markets. In *Proceedings of WWW*.

[19] Noah Golowich, Harikrishna Narasimhan, and David C Parkes. 2018. Deep Learning for Multi-Facility Location Mechanism Design.. In *Proceedings of IJCAI*.

[20] Negin Golrezaei, Adel Javanmard, and Vahab Mirrokni. 2019. Dynamic incentive-aware learning: Robust pricing in contextual auctions. In *Proceedings of NeurIPS*.

[21] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. 2018. Making the most of your samples. In *SIAM Journal on Computing*.

[22] Yash Kanoria and Hamid Nazerzadeh. 2014. Dynamic Reserve Prices for Repeated Auctions: Learning from Bids. In *Proceedings of WINE*.

[23] Alejandro M Manelli and Daniel R Vincent. 2007. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic theory* (2007).

[24] Andrés Muñoz Medina and Sergei Vassilvitskii. 2017. Revenue optimization with approximate bid predictions. In *Proceedings of NIPS*.

[25] Mehryar Mohri and Andres Munoz. 2015. Revenue optimization against strategic buyers. In *Proceedings of NIPS*.

[26] Jamie H Morgenstern and Tim Roughgarden. 2015. On the pseudo-dimension of nearly optimal auctions. In *Proceedings of NIPS*.

[27] R. B. Myerson. 1981. Optimal Auction Design. In *Math. Oper. Res.*, Vol. 6.

[28] Thomas Nedelec, Marc Abeille, Clément Calauzènes, Noureddine El Karoui, Benjamin Heymann, and Vianney Perchet. 2018. Thresholding the virtual value: a simple method to increase welfare and lower reserve prices in online auction systems. *arXiv preprint arXiv:1808.06979* (2018).

[29] Thomas Nedelec, Noureddine El Karoui, and Vianney Perchet. 2019. Learning to bid in revenue-maximizing auctions. *Proceedings of ICML* (2019).

[30] M. Ostrovsky and M. Schwarz. 2011. Reserve prices in internet advertising auctions: A field experiment. In *Proceedings of EC*.

[31] Renato Paes Leme, Martin Pal, and Sergei Vassilvitskii. 2016. A field guide to personalized reserve prices. In *Proceedings of WWW*.

[32] Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z Berkay Celik, and Ananthram Swami. 2017. Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security*.

[33] Weiran Shen, Sébastien Lahaie, and Renato Paes Leme. 2019. Learning to Clear the Market. In *Proceeding of ICML*.

[34] Weiran Shen, Pingzhong Tang, and Song Zuo. 2019. Automated mechanism design via neural networks. In *Proceedings of AAMAS*.

[35] Pingzhong Tang and Yulong Zeng. 2018. The price of prior dependence in auctions. In *Proceedings of EC*.

[36] Daan Wierstra, Tom Schaul, Jan Peters, and Juergen Schmidhuber. 2008. Natural evolution strategies. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*. IEEE, 3381–3387.

[37] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* 8, 3-4 (1992), 229–256.

[38] Andrew Chi-Chih Yao. 2017. Dominant-strategy versus bayesian multi-item auctions: Maximum revenue determination and comparison. In *Proceedings of EC*.

[39] Hanrui Zhang, Yu Cheng, and Vincent Conitzer. 2019. When Samples Are Strategically Selected. In *Proceedings of ICML*.