EC 2026 Paper #1811 Reviews and Comments
===
Paper #1811 Is Online Linear Optimization Sufficient for Strategic
Robustness?

Review #1811A
===
* Updated: May 6, 2026

Overall merit
----
6. Weak accept: the paper is competitive for EC, and I hope it can be
   accepted. I intend to advocate for it, but it should be considered
   against this year's competition. I expect it to be in the *top 25% of
   submissions*.

Paper summary

This paper studies strategically robust (to the seller's strategic reserve price setting) no-regret bidding in repeated Bayesian first-price auctions. Building on the concave quantile-space formulation of Kumar et al. (2024), the authors first show that seller revenue can be bounded by the Myerson benchmark plus the buyer’s linearized regret against the all-zero bidding strategy. This yields the paper’s main conceptual contribution: a black-box reduction showing that any online linear optimization (OLO) algorithm with sublinear linearized regret induces a strategically robust bidding algorithm. To obtain sharper guarantees, the paper then reparameterizes the original quantile-space action set into a probability simplex via bid-probability differences. This more structured space enables the use of standard OLO algorithms and improves the dependence on the bid discretization size from the prior $O(\sqrt{TK})$-type bound to an $O(\sqrt{T\log K})$-type bound in the known-distribution setting. The paper also extends the approach to unknown value distributions via a dominated empirical-distribution estimator, obtaining high-probability regret and robustness guarantees; in particular, this replaces the prior $O(\bar f^{1/2}\sqrt{TK})$-type guarantee with an $O(\sqrt{T}(\log K + \log(T/\delta)))$-type bound while also removing the bounded-density assumption from prior work.

Evaluation

Overall Evaluation

Overall, I found the paper very well written and easy to follow. Its main contribution, in my view, is a clean black-box reduction from OLO to strategically robust bidding, built on the observation that strategic robustness can be controlled through linearized regret. This observation, together with the example showing that ordinary regret is insufficient, is neat and insightful. The paper then adds a simple and useful reparameterization of Kumar et al.’s quantile-space formulation, moving from a strategy polytope to a simplex of bid probabilities. Although this step is conceptually natural, it proves quite useful in enabling sharper standard online-learning bounds and significantly improving the dependence on the bid discretization size. Taken together, these ingredients yield a clean and technically solid contribution, which is further strengthened by the extension to the unknown-distribution setting through a distribution-estimation step that also improves over prior work there.

At the same time, I am left with some broader modeling questions. In particular, the additive robustness benchmark appears closely tied to the fixed value-distribution assumption and the specific first-price auction structure, where the same concave utility representation governs both regret and seller revenue and leads to a $T \cdot \mathrm{Myer}(F)$ plus robustness-term benchmark. This makes the benchmark feel somewhat specific to the present formulation, even though it is well motivated within that setting. With that perspective in mind, I summarize the main strengths and weaknesses below.

Strengths

1. The paper is very clearly written and easy to follow, with smooth exposition and a well-structured presentation of the technical ideas.

2. The paper provides a neat black-box reduction from OLO to strategically robust bidding via linearized regret, and combines this with a simple simplex reparameterization that enables sharper regret bounds and a clean extension to the unknown-distribution setting.

Weaknesses

1. While the black-box reduction is conceptually interesting, the connection between strategic robustness and linearized regret appears somewhat incremental in isolation, as it builds directly on the concave quantile-space formulation of Kumar et al.

2. The simplex reformulation is clean and useful, though it reads more as a natural and well-motivated reparameterization than as a fundamentally new conceptual ingredient. In particular, the move from cumulative quantiles to bid probabilities appears relatively direct given the existing formulation.

3. The robustness analysis is structurally tied to comparison with the all-zero bidding strategy, but the final guarantees are obtained via generic worst-case OLO regret bounds over the entire simplex. It is unclear how much slack this introduces, and it would be interesting to know whether tighter direct bounds against that special comparator are possible.

Post Rebuttal Evaluation:

Thanks to the authors for the thoughtful responses! I found many of them compelling and appreciated the clarifications. After considering the discussion around these concerns and the helpful and strong points raised by the great reviewing team, I've decided to keep my evaluation as it was.

Questions for Authors

Questions for the Authors

Q1. The additive robustness benchmark appears closely tied to the fixed value-distribution assumption and to the particular first-price auction structure, where the same concave utility representation governs both regret and seller revenue. How would the framework change if valuations were drawn from time-varying distributions F_t, or even chosen adversarially, for example in a smoothly varying rather than i.i.d. manner? In such settings, would a benchmark such as $\sum_t \mathrm{Myer}(F_t)$, or perhaps a different notion of robustness, be more appropriate?

Q2. Given the close connection between the quantile-space formulation and the $T \cdot \mathrm{Myer}(F)$ benchmark, do the authors see alternative formulations or regret notions that could support stronger or qualitatively different robustness benchmarks? More broadly, what do the authors view as the most meaningful next directions for this line of work: further tightening the regret/robustness bounds, or developing more general notions of robustness beyond the current setting?

Q3. The structural analysis controls robustness through comparison with the all-zero bidding strategy, while the final guarantees rely on generic worst-case regret bounds over the full simplex. Is there hope for a tighter, more direct analysis against this particular comparator, rather than passing through worst-case regret?

Q4. The model assumes full feedback of the highest competing bid (or effective reserve) in each round. While this is standard in recent theoretical work, how sensitive are the results to this assumption? Do the authors see plausible extensions of the framework to weaker or partial-feedback settings?

ACM Policy Attestation

Yes

Forward to Journal

Yes

Review #1811B
===

Overall merit
----
7. Accept: the paper should certainly be accepted and I am comfortable
   making this judgment without knowing the competition. I expect it to be
   in the *top 10% of submissions*.

Paper summary

In this paper, the authors study repeated Bayesian first-price auctions. In this setting, a bidder—whose valuation for the item is drawn from a distribution—participates in a sequence of first-price auctions. In each round, the bidder observes the realization of her value, submits one of K equally spaced bids for the item, and then receives feedback in the form of the highest competing bid (or, alternatively, the reserve price), which reveals whether she has won the auction. The bidder may either know her own value distribution or have it unknown; both scenarios are considered in this work.

In this problem, it is highly desirable for the bidder's algorithm (mapping current value and past information to the bid to submit to the auction) to satisfy two key properties: first, it should be no-regret, that is its cumulative value over rounds is only a sublinear (in the number of rounds) additive factor away to that of the best fixed bidding strategy in hindsight. Second, the algorithm should be non-manipulable, that is a strategic seller would not be able to extract revenue larger than the Myerson's optimal mechanism (times the number of rounds) by more than a sublinear additive factor.

The authors present a blackbox reduction that takes an algorithm minimizing linearized regret and turns it into one that is additionally strategically robust. Specifically, their main contributions are described as follows:

* They first reformulate the problem in the quantile strategy space following Kumar et al. (2024), meaning that one cares about the probabilities of bidding above a certain value. They then prove that any algorithm with low linearized regret cannot be exploited by a strategic seller by more than the magnitude of the linearized regret itself. This stands in contrast with usual regret, since it is proven that there exists a distribution and an algorithm that, despite suffering sublinear (and even negative regret), could be heavily exploited by a strategic seller (even in the known distribution setting).

* They further redefine this concave formulation to only care about the probability of bidding a specific bid value and over the simplex. Then, in the known value distribution setting, their procedure takes a blackbox no-linearized-regret algorithm which outputs a probability distribution q_t and submits the bid corresponding to the quantile (of the value distribution calculated on q_t) the realized value lands in. The gradient of the utility at q_t given the highest competing bid is then forwarded to the blackbox algorithm. This yields sublinear regret as well as sublinear strategic robustness, where the reformulation over the simplex allows a tight dependence on the number of bids.

* Their main result lies in the unknown value distribution case. In this case, two steps of the previous procedure can no longer be carried out due to the absence of knowledge of the value distribution. First, the bidder cannot check the exact quantile the value lands in; second, the bidder cannot compute the gradient of the utility function exactly, since this depends on the underlying value distribution.

* Starting from the empirical distribution, they define the *Dominated Continuous Empirical Distribution* (DCED) which interpolates the empirical distribution into a continuous version to remove discretization errors. The DCED also makes a dominance adjustment, that is, it shifts a small probability mass to 0 so that the true distribution stochastically dominates, thereby preserving the revenue ordering. Crucially, this construction keeps, with high probability, errors independent of the number of bids and adds only a small regret term, again independent of the number of bids. In the end, this allows to essentially mirror the procedure carried out in the known distribution setting, but based on the DCED, and allows for the nearly-tight regret and strategic robustness rates.

Evaluation

I found this paper both very enjoyable to read and technically strong. The fact that such a general result can be established using a black-box linearized regret minimizer is, in itself, quite surprising. The overall structure of the analyses is not only clear but also elegant.

I will highlight below several points where I believe the paper's contributions are particularly crisp and insightful:

* First, the proof of Theorem 2 is elegant and, in retrospect, seems like the natural way to approach the problem.

* What I found most interesting is the construction of the *Dominated Continuous Empirical Distribution*—it is cleverly designed and the role of each step in the overall argument is easy to follow.

* Moreover, the paper's results substantially generalize the earlier work of Kumar et al. (2024), who first demonstrated that Online Gradient Ascent (OGA) can achieve both sublinear regret and sublinear strategic robustness—specifically, $O(\sqrt{TK})$ in the known-distribution case and $O(K\sqrt{\bar f T})$ in the unknown case, where $\bar f$ is the maximum density over the support of the value distribution.

* This paper improves upon that work in several ways. It shows that a broad class of algorithms can achieve the same guarantees without requiring a specialized analysis like that of OGA. Furthermore, the dependence on the problem's natural parameters T and K is tighter and nearly optimal. Importantly, the results also avoid reliance on the artificial max-density parameter.

As a minor note, on page 2, "subliear" should be corrected to "sublinear." Additionally, at the beginning of the proof of Lemma 5, I believe you intended to write $q_{t,0} = 0$, right?

Questions for Authors

Q1: The observation that low regret does not guarantee strategic robustness is quite intriguing. However, I was wondering about the following point: you show that there exists a distribution and an algorithm for which even low (or negative) regret does not imply strategic robustness. Yet, such an algorithm is likely to perform quite poorly in terms of regret for the worst possible distribution (I didn't actually find a distribution where this is the case, but I believe there should be one). This leads me to ask whether *low worst-case regret* would, in contrast, imply strategic robustness. To ensure this question is not already covered by the main results, I tried to think of an algorithm that achieves low worst-case regret (in the concave setting) but high worst-case *linearized* regret. I could not come up with any such algorithm, since most algorithms with low worst-case concave regret also exhibit low linearized regret—this is typically how sublinear regret is proven in the first place. I would be interested in hearing your thoughts on whether this distinction makes sense and whether such a counterexample could exist.

Q2: Regarding the analysis, it seems that knowing the highest competing bid is not strictly necessary to compute the utility, since a single bit of feedback—whether or not the bidder won—would suffice for that. However, the *value* of the highest competing bid appears crucial for computing gradients, regardless of whether the bidder knows the underlying distribution (please correct me if that is not the case). Do you think there might be a way to reconstruct or approximate gradient information using only binary feedback (whether or not the bidder won)? I am not sure how one would achieve this, but I would appreciate your perspective. Of course, feel free to disregard this if it falls outside the paper's intended scope.

ACM Policy Attestation

Yes

Forward to Journal

Yes

Review #1811C
==========

Overall merit
---
6. Weak accept: the paper is competitive for EC, and I hope it can be
   accepted. I intend to advocate for it, but it should be considered
   against this year's competition. I expect it to be in the *top 25% of
   submissions*.

Paper summary

The paper studies strategically robust learning in repeated first-price auctions. In each round, the buyer draws a value from a potentially unknown value distribution, and her strategy is a mapping from the value to a set of equally spaced bids between 0 and 1. Strategic robustness means that the seller cannot extract more than the Myerson revenue on average, plus some vanishing terms. The buyer also aims to have no regret when competing with the best mapping.

The key technical contribution of the paper is as follows. First, it provides a more general concave formulation for the discrete-time problem. The buyer's strategy is represented in quantile space, i.e., as a map from the quantile of her value to the discrete bids, and is further identified using a probability simplex. Second, given the concave formulation, it uses a linear approximation to upper bound the regret and, more importantly, proves that the seller's revenue cannot exceed the Myerson revenue plus the linear regret approximation. With these observations, the paper produces a reduction that turns any online linear optimization algorithm into a strategically robust algorithm by using the OLO algorithm to minimize the linearized regret.

By applying the result to MWU, the paper gives a strategically robust algorithm whose regret has optimal polylogarithmic dependence on the size of the bid space, which improves upon the conventional strategically robust swap-regret algorithm that has linear dependence on that size.

The paper also extends to the case where the value distribution is unknown and removes the requirement that the strategically robust algorithm needs the value distribution. It uses a similar strategy to that in the sample complexity of learning optimal auction papers to learn the value distribution from samples.

Evaluation

Strengths: I think the paper is an interesting generalization of [Kumar et al.] and provides further insights into the usefulness of quantile space. The optimal regret result derived from the reduction is also quite interesting. The presentation of the lemmas and theorems is also quite easy to understand.

Weaknesses: I do not find any major weaknesses. The paper does not have explicit results that extend beyond first-price auctions, which a concurrent work mentioned is able to do, so it seems to be weaker, but I think their reduction and algorithm are almost identical. The unknown value distribution case is not particularly technically interesting for those familiar with the technique, and I do not think there is anything significantly different from the offline case, but it is still a substantial improvement over [Kumar et al.].

Questions for Authors

Q1: My understanding is that your reduction and algorithm are identical to those in the concurrent work and can be easily generalized to other auction formats without significant changes in notation. Is this correct? (I am not suggesting any changes in this direction for your draft. I just want to confirm that my understanding is accurate.)

ACM Policy Attestation

Yes

Forward to Journal

Yes

Review #1811D
==========

Overall merit
---
4. Weak Reject: the paper is competitive for EC, but should be rejected. I would not argue strongly against acceptance, but another PC member would need to convince me to support acceptance.

Paper summary

The paper studies the problem of learning to bid in repeated first-price auctions. In particular, a single bidder repeatedly participates in sequential first-price auctions where her values are i.i.d., the bid space is finite, and the reserve price is chosen adversarially. Importantly, the reserve price is revealed after the auction. The goal is to design an online learning algorithm that can use this feedback to learn good bidding strategies (maps from values to bids) over time. As is typical in online learning, the authors use (pseudo) regret against the best static strategy to measure the rate of learning. Going beyond the typical objective of regret, the authors also consider strategic robustness as analyzed in Braverman et al. [2018] and Kumar et al. [2024]: an online learning algorithm is said to be strategically robust if no sequence of reserve prices can extract more revenue from it than simply posting the optimal Myerson reserve price in every auction.

The paper builds on the concave formulation of Kumar et al. [2024], which reformulated the problem of optimizing the bidding strategy (a map from values to bids) in quantile space (the probability of bidding a certain value under monotonic strategies) and showed that this reformulation results in a concave optimization problem. Therefore, this reformulation yields an OCO problem, and any no-regret algorithm for OCO can be used to obtain regret guarantees. Kumar et al. [2024] also showed that Online Gradient Ascent, which is no-regret for OCO, is additionally strategically robust. The current paper shows that this fact is true not just for OGA but for any OCO algorithm with low linearized regret applied to the quantile reformulation. It does so by showing that the proof technique of Kumar et al. [2024] goes beyond OGA and applies more broadly to any algorithm with low linearized regret. Finally, the authors extend their result to the case where the value distribution is not known a priori and thus cannot be used to obtain the gradients in quantile space. They show that an empirical estimate of the value distribution, perturbed to stochastically dominate the true CDF with high probability, can be used to estimate these gradients and ensure strategic robustness in the unknown-value-distribution setting.

Evaluation

Strengths:
The paper is well written, with a clear exposition. Its main finding, that the concave formulation in quantile space (Kumar et al. [2024]) is what drives strategic robustness in the single-buyer setting, is important. In particular, it is not only OGA that is robust, but any no-regret algorithm with low linearized regret---even FTRL or any other mean-based algorithm---applied in quantile space is also robust. This is important because mean-based algorithms used directly to select bids for each value are known to be exploitable (Braverman et al. [2018]).

Weaknesses:
Although it is good to know that the proof technique of Kumar et al. [2024] extends to all low-linearized-regret algorithms, and that such algorithms are strategically robust when applied to the concave formulation, the results are limited to the single-bidder setting and do not address incentive compatibility for the bidder. Kumar et al. [2024] show that the bidder does not regret truthfully reporting her valuations to the OGA algorithm in hindsight. They then use this fact to show that the seller cannot extract more average revenue from a collection of buyers employing OGA than the Myerson optimum. Thus, although the current paper provides an extension of the results and proofs of Kumar et al. [2024] in the single-bidder setting, the extension is somewhat limited and does not introduce substantial new techniques in my view. Additionally, while the improved dependence on K that comes with the Hedge algorithm is good to have, it comes at the cost of higher regret ($O(\sqrt{T})$ vs. $O(\log T)$ for OGA) when the competing bids are stochastic.

[Optional] Feedback for Authors

Page 7: Please clarify how the non-overbidding condition make the strategy set depend on F. I would

assume it is always possible to bid the minimum of the value and the bid recommended by the strategy.

ACM Policy Attestation

Yes

Forward to Journal

Yes

ConcreteQuestions Response by Weiqiang Zheng <weiqiang.zheng@yale.edu> (Author) (574 words)

RAQ1: In the known value setting, we can generalize our results to time-varying distributions (as our per-step proof still holds) with the proposed benchmark. In the unknown value setting, estimating the value no longer works for time-varying distributions, so our results do not generalize. The setting of adversarially chosen values is harder and beyond the scope of this paper since the concave reformulation relies on the value distribution being continuous. We note that the literature on strategic robustness also focuses on repeated fixed games.

RAQ2: A qualitatively stronger benchmark seems unlikely as the seller can always use Myerson's auction to get $T\cdot\mathrm{Myer}(F)$ (assuming the buyer is rational and learns to best respond). But we remark that in the proof we obtain a fine-grained instance-dependent bound of the revenue $\sum_t p_{t,i}\cdot{F}^-(1- p_{t,i})$, which is stronger than $T\cdot\mathrm{Myer}(F)$ (see also remark 2 in [KSS24]).

We think the most important next direction is to characterize the class of games in which simple learning algorithms also ensure strategic robustness

RAQ3: This is an insightful question. Indeed, we can apply the AB-prod algorithm [Sani-Neu-Lazaric'14] to obtain better bounds on the strategic robustness (they proved for standard regret, but their algorithm and analysis can be generalized to linearized regret). Specifically, we choose A to be the Hedge algorithm and B to be the algorithm that always chooses the all-zero strategy (which has zero regret against that comparator). Then the AB-prod algorithm guarantees that the worst-case regret is $O(\sqrt{T\log K + T\log T})$, and the regret against the all-zero strategy is $O(1)$, leading to constant strategic robustness. We will add these corollaries in the revised version of the paper.

RAQ4: We can consider two weaker feedback settings. The first one is binary win-lose feedback without value; the extension is not immediate and requires new insights. The second is that the bidder sees the highest competing bid only when he wins/loses the auction; in this case, we can explore with a small ϵ probability (bidding 0 or 1) to get the value and obtain an unbiased estimator of the gradient. The caveat is that the regret bound would be worse than $\sqrt{T}$ due to the estimation variance.

RBQ1: The distinction makes sense, and an algorithm that has low worst-case regret but high worst-case linearized regret does exist. A simple construction would be as follows: the algorithm first follows the sequence of strategies in Theorem 3 that gives negative regret for the specific distribution, and switches to a generic no-regret algorithm when it detects its regret is $\Omega(\sqrt{T})$. So its worst-case regret is sublinear; however, its worst-case linearized regret is linear when the distribution is as in Theorem 3 and is not strategically robust. This is analogous to the construction that no-regret learning dynamics can converge to any particular CCE (they first play according to that CCE and switch to a no-regret algorithm upon detecting a deviation). As a result, low worst-case regret does not imply strategic robustness.

RBQ2: See also R1Q4. In our case, knowing the highest competing bid is necessary even to compute the utility of a bidding strategy (which is a mapping from value to bids) since it requires computing the utility of every possible bid.

RCQ1: Yes, our reduction and algorithm are identical to the concurrent work and can be generalized to other auction formats. To extend to other auction formats, the change is in the gradient formula, which depends on the allocation and payment rule of that auction.

FactualCorrections Response by Weiqiang Zheng <weiqiang.zheng@yale.edu> (Author) (116 words)

>RD: "Thus, although the current paper provides an extension of the results and proofs of Kumar et al. [2024] in the single-bidder setting, the extension is somewhat limited and does not introduce substantial new techniques in my view."

We believe our results help demystify OGD's success by showing that the key property is not OGD per se, but rather the minimization of linearized regret within the concave formulation. This is not established in [Kumar et al., 2024].

In the unknown-distribution setting, we introduce a new technique based on the dominated empirical distribution that overcomes the limitations in [Kumar et al., 2024]. To the best of our knowledge, this technique is new in the strategic robustness literature.

Review #1811E
==========

Overall merit
---
6. Weak accept: the paper is competitive for EC, and I hope it can be accepted. I intend to advocate for it, but it should be considered against this year's competition. I expect it to be in the *top 25% of submissions*.

Paper summary

The paper studies repeated Bayesian first-price auctions with one buyer. In each round, the buyer observes an i.i.d. value from a distribution F, bids from a finite grid, then observes the highest competing bid or effective reserve. The buyer wants two guarantees: low regret against the best fixed value-to-bid mapping, and strategic robustness, meaning that a strategic seller cannot extract cumulative revenue more than T·Myer(F) plus a sublinear term. The paper builds directly on the concave quantile-space formulation of Kumar, Schneider, and Sivan, which derived similar results for online gradient ascent. This paper makes the crisp conceptual contribution that one of the important ingredients in the Kumar et al. argument is not online gradient ascent specifically, but the fact that it minimizes linearized regret --- i.e. it gets no regret not just to the convex loss functions it observes but to the linear loss functions derived from their gradients.

The main technical observation is Theorem 2. For a monotone bidding strategy represented by quantile variables p, the seller’s revenue in a round can be algebraically related to the inner product of the utility gradient with p. This connection lets them plug in arbitrary algorithms with linearized regret guarantees. For example, plugging in multiplicative weights yields an O(sqrt{TlogK}) regret and robustness bound in the known-F setting, improving the earlier O(sqrt{TK})-dependence that comes from online gradient ascent. They then give an extension to the unknown F setting in which they must learn it online.

Evaluation

The strength of this paper is that it identifies linearized regret as the right abstraction rather than other properties specific to online gradient descent, which was unclear from prior work. Not only is this conceptually important for our understanding of what is going on, it opens the door to improved bounds by plugging in multiplicative weights to their black box linearized regret reduction.

The weaknesses of this paper are its limited scope: it restricts attention to one bidder first price auctions with iid values and full feedback. The technical novelty on top of Kumar et al. is also limited once one has the crucial insight that linearized regret is the key quantity of interest.

Despite these weaknesses, I think that this paper clears the bar for EC because of its crisp, clarifying insight. The paper is also clearly written.

ACM Policy Attestation

Yes

Forward to Journal

Yes