2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)

Benchmark Design and Prior-independent Optimization

Jason Hartline, Aleck Johnsen, Yingkai Li
Department of Computer Science
Northwestern University
Evanston, IL, United States
hartline@u.northwestern.edu, aleckjohnsen@u.northwestern.edu, yingkai.li@u.northwestern.edu

***Abstract*—This paper compares two leading approaches for robust optimization in the models of online algorithms and mechanism design. Competitive analysis compares the performance of an online algorithm to an offline benchmark in worst-case over inputs, and prior-independent mechanism design compares the expected performance of a mechanism on an unknown distribution (of inputs, i.e., agent values) to the optimal mechanism for the distribution in worst case over distributions. For competitive analysis, a critical concern is the choice of benchmark. This paper gives a method for selecting a good benchmark. We show that optimal algorithm/mechanism for the optimal benchmark is equal to the prior-independent optimal algorithm/mechanism.**

We solve a central open question in prior-independent mechanism design, namely we identify the prior-independent revenue-optimal mechanism for selling a single item to two agents with i.i.d. and regularly distributed values. We use this solution to solve the corresponding benchmark design problem. Via this solution and the above equivalence of prior-independent mechanism design and competitive analysis (a.k.a. prior-free mechanism design) we show that the standard method for lower bounds of prior-free mechanisms is not generally tight for the benchmark design program.[1]

I. Introduction

There are two leading approaches for robust optimization in the models of mechanism design and online algorithms (see detailed historical context deferred to the end of this section). A key property of these environments is that with the given constraints (incentive compatibility or on-line arrivals) the best outcome is unachievable pointwise. The first approach considers a benchmark and looks for a mechanism (resp. algorithm) that pointwise approximates the benchmark, a.k.a., prior-free approximation (resp. competitive analysis). The choice of benchmark is an important variable of this approach. The second approach assumes that the input is drawn at random from an unknown distribution in a family and looks for a mechanism that approximates, in worst-case over distributions in the family, the performance of the Bayesian optimal mechanism for the distribution, a.k.a., prior-independent approximation. (Respectively, this approach could also be applied to online algorithms.) This paper formalizes the problem of designing a good benchmark for prior-free approximation (resp. competitive analysis) and connects this benchmark design problem to prior-independent optimization.

[1]For the full version of this work, see https://arxiv.org/abs/2001.10157.

Terminology and concrete examples from mechanism design are henceforth adopted for the majority of the paper. However, the main result relating benchmark design to prior-independent optimization does not rely on any specifics from mechanism design. Instead it applies to families of mechanisms represented by a family of functions from inputs to performances. In Section V we instantiate the framework for online learning and give results that parallel the specific developments for mechanism design.

The choice of benchmark in the first approach impacts the ability of approximation with respect to the benchmark to distinguish between good and bad mechanisms. On one hand a benchmark should be an upper bound on what is achievable by a mechanism; otherwise, approximating it does not necessarily mean that a mechanism is good. On the other hand it should not be too loose an upper bound; otherwise, neither good nor bad mechanisms can obtain good approximations and the degree to which we can distinguish good and bad mechanisms via the benchmark is limited.

As illustration, consider comparing the revenue of mechanisms for selling a digital good to n agents with values $\boldsymbol{v} = (v_1, \ldots, v_n)$ bounded on $[1, H]$ to one of two benchmarks, the *sum-of-values* benchmark $\sum_i v_i$ and the *price-posting-revenue* benchmark $\max_i i\, v_{(i)}$ where $v_{(i)}$ is the ith highest value. To ensure benchmarks give an upper bound on revenue, Hartline and Roughgarden (2008) suggested that the benchmark satisfies the property that, for any distribution on inputs from a given family of distributions, the expected benchmark (over the same distribution of inputs) be at least the expected performance of the optimal mechanism that knows the distribution. Thus approximation of the benchmark implies approximation of the optimal mechanism for any of the distributions in the family. We refer to this constraint as *normalization*. Both the sum-of-values and price-posting-revenue benchmarks are normalized (Hartline and Roughgarden, 2008).

Not all normalized benchmarks are equally good at discriminating between good and bad mechanisms. Consider the following two mechanisms. The *random-sampling* mechanism partitions the agents at random and offers the optimal price from each part to the other part (Goldberg et al., 2006).

2575-8454/20/$31.00 ©2020 IEEE
DOI 10.1109/FOCS46700.2020.00036

The *random-power-pricing* mechanism posts a take-it-or-leave-it price drawn from the uniform distribution on powers of two in $[1, H]$ (Goldberg and Hartline, 2003). These mechanisms and the benchmarks of the preceding paragraph are related as follows. On all inputs $\boldsymbol{v}$, the sum-of-values benchmark exceeds the random-power-pricing mechanism by a $\Theta(\log H)$ factor. On all inputs $\boldsymbol{v}$, the random sampling mechanism and the price-posting-revenue benchmark are $\Theta(1)$. Moreover, the performance of the latter benchmark and mechanism are always sandwiched between the former benchmark and mechanism and can equal either of them up to $\Theta(1)$. From this analysis, we see that the loose benchmark of sum-of-values does not discriminate between good mechanisms like random-sampling and bad mechanisms like random-power-pricing.

The preceding discussion suggests a benchmark design problem of identifying the normalized benchmark to which the tightest approximation is possible. We refer to the tightest approximation possible for a benchmark as its *resolution* (see Definition 3), as up to this factor the benchmark cannot distinguish between good and bad mechanisms. The sum-of-values benchmark has logarithmic resolution while the price-posting-revenue benchmark has constant resolution.

In summary: A family of distributions over inputs induces a class of normalized benchmarks. The benchmark admitting the tightest approximation, i.e. having smallest resolution, is optimal. The mechanism achieving this approximation is the prior-free optimal mechanism for the benchmark. Normalization of the benchmark implies that the prior-free approximation factor of a mechanism (to the benchmark) is at least its prior-independent approximation factor (for the normalizing family of distributions). A natural question is how this prior-free approach, and its optimal mechanism, compares to directly identifying the prior-independent optimal mechanism, i.e., the one with the best worst-case over distributions approximation to the Bayesian optimal mechanism.

The first main result of the paper is the general result that optimal benchmark design is equivalent to prior-independent optimization (Section II). Described in the context of mechanism design, this result shows that the prior-free optimal mechanism for the optimal benchmark is the prior-independent optimal mechanism and that the optimal benchmark is simply the prior-independent optimal mechanism scaled up by its approximation factor (so as to satisfy the normalization constraint).[2] Consequently, it is not possible to identify optimal benchmarks and their corresponding optimal mechanisms for problems in which we are unable to solve the prior-independent optimization problem.

Our second main result is to solve the benchmark optimization problem (equivalently: solve the prior-independent mechanism design problem) for the problem of maximizing revenue from the sale of a single item to two agents for the family of i.i.d. regular value distributions (Section III). This result answers a major question left open from Dhangwatnotai et al. (2015), Fu et al. (2015), and Allouah and Besbes (2018). The optimal mechanism is a mixture between the second-price auction, where each agent is offered a price equal to the highest of the other agents' values, and the auction where these prices are scaled up by a factor of about 2.5. Our solution to this central open question is the first example of a prior-independent optimal mechanism that arises as the solution to a non-trivial optimization problem and is not a standard mechanism from the literature (though the mechanism does fall into the lookahead family of mechanisms described by Ronen (2001) and has the same form as mechanisms used to prove bounds in Fu et al. (2015) and Allouah and Besbes (2018)). Our construction pins down the worst-case family of distributions. To solve the equivalent of benchmark design and prior-independent optimization, the approach of this paper is to directly solve the prior-independent optimization problem and use that solution to solve the benchmark design problem.

A key component of optimal benchmark design is tight lower bounds on the prior-free approximation of a benchmark. There is a standard method for lower bounding the approximation ratio of the best mechanism for a given benchmark (Goldberg et al., 2006). Consider the distribution over inputs for which all mechanisms achieve the same performance, e.g., for revenue maximization in mechanism design this distribution is the so-called *equal-revenue* distribution. The ratio of the expected value of the benchmark on this distribution to the revenue of any mechanism (all "undominated" mechanisms are the same) gives a lower bound on the approximation factor of any mechanism to the benchmark. Hartline and McGrew (2005) proved that this approach is tight for a large family of benchmarks and $n = 3$ agents. Chen et al. (2014) proved that this approach is tight for a large family of benchmarks and a general number n of agents. Thus, a natural approach to solve the benchmark optimization problem is to relax the program to optimize, not the approximation factor of the best mechanism, but the lower bound on the best approximation factor from the above approach. Our third main result – without being able to explicitly solve this relaxed program – is that the relaxation is not without loss, i.e., it gives a normalized benchmark for which the lower bound is not achievable by any mechanism (Section IV).

Our proof of the third main result follows from the second main result and a counter example that shows that there is a benchmark that achieves a lower objective value (for the

[2]The analysis that proves this equivalence is straightforward and, perhaps, obvious in hindsight. As we will describe below, there is an alternative benchmark design program which we view, in hindsight, as a relaxation of our benchmark design program. The prior literature strongly suggested that this alternative program and our program are equivalent; our third result shows that in fact the relaxation is lossy.

relaxed benchmark program) than the approximation ratio of the optimal prior-independent auction (which our first result shows to be the optimal objective value of the original benchmark program).

Our second and third results are proved under the restriction of mechanisms that are (a) dominant strategy incentive compatible (DSIC), i.e., where truthtelling is a good strategy for each agent regardless of the strategies of other agents, and (b) scale invariant. It is known that there are environments for prior-independent mechanism design where DSIC is not without loss (Feng and Hartline, 2018). The restricted family of DSIC mechanisms is interesting even if it is with loss; however, far more study of non-incentive-compatible mechanisms is warranted. It is not known whether scale invariance is without loss or not, though Allouah and Besbes (2018) conjecture that it is without loss. This question is important and remains open.

There is an important negative interpretation of our first result, that the prior-free benchmark optimization problem and the prior-independent optimization problem give the same answer (which has consequences for both mechanism design and online algorithms). One reason to prefer prior-free analysis over prior-independent analysis is that it could be more robust. Of course to be more robust, as Hartline and Roughgarden (2008) have recommended, we need to choose a normalized benchmark, i.e., one for which prior-free approximation implies prior-independent approximation. With many possible normalized benchmarks, we need a method for selecting one. We have adopted a natural formal method for selecting one, namely, the one that admits the tightest approximation is the best one. However, the results of the paper show that the answer we will then get from studying prior-free approximation of this optimal benchmark is the same as the answer we will get from the original prior-independent question. Thus, there is no added robustness from the prior-free approximation of the optimal benchmark (over prior-independent analysis). Of course, there are environments where increased robustness can informally be observed from prior-free approximation of ad hoc, i.e., non-optimal, benchmarks. This observations suggests the main open question of this paper which is to identify a rigorous framework for evaluating prior-free benchmarks which leads to an additional desired robustness over the prior-independent framework. We will formally describe this shortcoming of the framework with the example environment of no-regret online learning in Section V.

Historical Context of Online Algorithms and Mechanism Design: Online algorithms have been analyzed via a worst-case competitive analysis since Sleator and Tarjan (1985) with textbooks on the subject, e.g., Borodin and El-Yaniv (2005). In competitive analysis, the performance of an online algorithm is measured as its worst case ratio to the optimal offline algorithm. For some problems this measure is too pessimistic, occurring when no good ratio is achievable by any algorithm and therefore good algorithms are not meaningfully separated from bad algorithms. The two approaches for resolving this issue are to either (a) restrict the offline algorithm to which the performance of the online algorithm is compared or (b) restrict the family of inputs that are considered. For this paper, the most relevant example of (a) comes from online learning where a learning algorithm's regret is measured with respect to the best fixed action in hindsight, i.e., to the optimal offline algorithm that is restricted to choose the same action in each time period (Littlestone and Warmuth, 1994; Freund and Schapire, 1997). The most relevant example of (b) for this paper is the diffuse adversary model of Koutsoupias and Papadimitriou (2000) which evaluates an algorithm as the ratio between its expected performance and the optimal offline performance in worst-case over a family of distributions on inputs.

Competitive analysis was introduced to the design of mechanisms by Goldberg et al. (2006). Hartline and Roughgarden (2008) revisited the choice of benchmark of Goldberg et al. (2006) and identified the normalization constraint. For auction settings, Devanur et al. (2015) give a simpler normalized benchmark based on relaxing the incentive constraints to constraints of envy-freedom. The prior-independent corollary of prior-free approximation of the benchmarks of Hartline and Roughgarden (2008) motivated the consideration of relaxing the assumption of worst-case inputs in a similar fashion to approach (b) above. Dhangwatnotai et al. (2015) considered prior-independent mechanism design as a first-order goal and since then it has been the subject of a flourishing area of research. For revenue maximization in the sale of an item to one of two agents with values drawn from an i.i.d. regular distribution, Dhangwatnotai et al. (2015) show that the second price auction is a 2-approximation. Fu et al. (2015) gave a randomized mechanism showing that this factor of 2 is not tight. Upper and lower bounds on this canonical problem were improved by Allouah and Besbes (2018) to be within $[1.80, 1.95]$.[3] For this two agent problem with i.i.d. values from a distribution in the subset of regular distributions that further satisfy a monotone hazard rate condition, Allouah and Besbes (2018) show that the second-price auction is optimal. See the full version for more-detailed historical context.

II. Benchmark Optimization Is Prior-independent Optimization

We formulate the benchmark optimization problem in abstract terms and prove that it is equivalent to prior-independent optimization. Our framework and results in this section hold generally for algorithm design, however, we will adopt notation and terminology for mechanism design to maintain consistency with the discussion in subsequent

[3]Their lower bound of 1.80 holds under the additional assumption of scale invariance.

sections. Notably, the development of this section makes no assumptions on the families of distributions over the input that are considered.

Denote the space of inputs by $\mathcal{V}$ and an input in this space by $\boldsymbol{v}$. Denote a family of distributions over input space by $\mathcal{F} \subset \Delta(\mathcal{V})$ and a distribution in the family by F. Denote a family of feasible mechanisms by $\mathcal{M}$ and a mechanism in this family by M. Denote a family of benchmarks by $\mathcal{B}$ and a benchmark in this family by B. For our purposes we will view both a mechanism and a benchmark as a function that maps the input space to an expected performance (e.g., in the case the mechanism is randomized), denoted respectively by $M(\boldsymbol{v})$ and $B(\boldsymbol{v})$. When evaluating the performance of a mechanism or benchmark in expectation over the distribution we adopt the short-hand notation $M(F) = \mathbf{E}_{\boldsymbol{v}\sim F}[M(\boldsymbol{v})]$ and $B(F) = \mathbf{E}_{\boldsymbol{v}\sim F}[B(\boldsymbol{v})]$.

In these abstract terms we formally define the Bayesian, prior-independent, and prior-free optimization problems.

Definition 1. *The* Bayesian optimal mechanism design *problem is given by a distribution F and family of mechanisms $\mathcal{M}$ and asks for the mechanism OPT_F with the maximum expected performance:*

$$\mathrm{OPT}_F = \operatorname*{argmax}_{M\in\mathcal{M}} M(F). \qquad (\mathrm{OPT}_F)$$

Definition 2. *The* prior-independent mechanism design *problem is given by a family of mechanisms $\mathcal{M}$ and a family of distributions $\mathcal{F}$ and solves the program*

$$\beta = \min_{M\in\mathcal{M}} \max_{F\in\mathcal{F}} \frac{\mathrm{OPT}_F(F)}{M(F)}. \qquad (\beta)$$

Next, ρ^B is the approximation ratio of the optimal prior-free mechanism for benchmark B; in the subsequent discussion of benchmark design, we reinterpret ρ^B as the *resolution* of benchmark B.

Definition 3. *The* prior-free mechanism design *problem is given by a family of mechanisms $\mathcal{M}$ and a benchmark B and solves the program*

$$\rho^B = \min_{M\in\mathcal{M}} \max_{\boldsymbol{v}\in\mathcal{V}} \frac{B(\boldsymbol{v})}{M(\boldsymbol{v})}. \qquad (\rho^B)$$

Recall from the introduction, benchmarks with small resolution are better at differentiating good mechanisms from bad one. Both prior-independent and prior-free mechanism design problems are searching for mechanisms with robust performance guarantees. In principle, prior-free guarantees can provide more robustness than prior-independent guarantees as the guarantee is required to hold pointwise on all inputs rather than in expectation according to the distribution. Whether or not a prior-free guarantee is meaningful depends on the choice of benchmark. The possibility that some benchmarks might be better than others for meaningfully quantifying the performance of a mechanism suggests that benchmarks themselves can be optimized.

Hartline and Roughgarden (2008) recommend restricting attention to benchmarks that satisfy the following normalization property which requires that the benchmark is an upper bound on the optimal performance of a mechanism. Rather than measure this optimal performance pointwise as is common with online algorithms, Hartline and Roughgarden (2008) recommend measuring this optimal performance in expectation with respect to any distribution in a family of distributions. This way of measuring the optimal performance can take into account the constraints on the mechanism, i.e., that $M \in \mathcal{M}$.[4] The normalization constraint implies a strong guarantee: A mechanism that is a ρ approximation to a normalized benchmark guarantees a ρ prior-independent approximation.

Definition 4. *A benchmark B is* normalized *for a family of distributions $\mathcal{F}$ and family of mechanisms if for every distribution in the family the expected benchmark is at least the optimal expected performance, i.e.,*

$$B(F) \geq \mathrm{OPT}_F(F), \quad \forall F \in \mathcal{F}.$$

Denote the normalized benchmarks for $\mathcal{F}$ by $\mathcal{B}(\mathcal{F})$.

Proposition 1 (Hartline and Roughgarden, 2008)**.** *If mechanism M is a prior-free ρ approximation of a benchmark B normalized to distributions $\mathcal{F}$ then its prior-independent approximation for distributions $\mathcal{F}$ is at most ρ.*

The point of mechanism design is a principled method for choosing one mechanism over another. The prior-free framework described above gives such a method only in so far as approximation of the benchmark distinguished between good mechanisms and bad ones. For example, no mechanism will approximate a prior-free benchmark that is too large, thus, good mechanisms will not necessarily be separated from bad ones. Recall the discussion of the sum-of-values and posted-price-revenue benchmarks in the introduction. One way to quantify the inability of a benchmark to discriminate is by considering the approximation factor of the best mechanisms for the benchmark, i.e., the benchmark's resolution (Definition 3). Our benchmark design program aims to identify the benchmark with the finest resolution.

Definition 5. *The* benchmark design *problem for family of distributions $\mathcal{F}$, family of mechanisms $\mathcal{M}$, and space of inputs $\mathcal{V}$ solves the program*

$$\gamma = \min_{B\in\mathcal{B}(\mathcal{F})} \rho^B = \min_{B\in\mathcal{B}(\mathcal{F})} \min_{M\in\mathcal{M}} \max_{\boldsymbol{v}\in\mathcal{V}} \frac{B(\boldsymbol{v})}{M(\boldsymbol{v})}. \qquad (\gamma)$$

We are now ready to state and prove the main result of this section, that the benchmark design problem and the prior-independent mechanism design problem are equivalent.

[4]For mechanism design, these constraints will be incentive compatibility and individual rationality. For online algorithms these constraints are that the current decision must be made before the future input is known.

Before doing so it should be noted that it is not generally understood how to solve these problems. (We will, however, give a solution to a paradigmatic prior-independent mechanism design problem in the next section.)

Theorem 1. *For any family of distributions $\mathcal{F}$, any set of mechanisms $\mathcal{M}$, benchmark design is equivalent to prior-independent mechanism design, i.e., $\gamma = \beta$, and the optimal benchmark is given by the performance of the prior-independent optimal mechanism scaled up by its approximation ratio β.*

This theorem follows from a corollary of Proposition 1 which shows that $\beta \leq \gamma$ and the following lemma which shows that $\gamma \leq \beta$.

Corollary 1. *For any families of distributions and mechanisms, the prior-independent optimal ratio β is at most the optimal benchmark ratio γ, i.e., $\beta \leq \gamma$.*

Proof: By the definition of program (γ), the optimal mechanism for the optimal benchmark is a prior-free γ approximation. By Proposition 1 and the normalization of the optimal benchmark, this mechanism is a prior-independent γ approximation. The optimal prior-independent mechanism is no worse, i.e., $\beta \leq \gamma$. ■

Lemma 1. *For any families of distributions and mechanisms, the optimal benchmark ratio γ is at most the prior-independent optimal ratio β, i.e., $\gamma \leq \beta$.*

Proof: Consider the prior-independent optimal mechanism M^* with approximation β. Define the benchmark

$$B^*(\boldsymbol{v}) = \beta\, M^*(\boldsymbol{v}), \tag{1}$$

i.e., the benchmark is the performance of the prior-independent optimal mechanism scaled up by its approximation factor. Taking the expectation of $\boldsymbol{v}$ drawn from any distribution F, we have

$$B^*(F) = \beta\, M^*(F). \tag{2}$$

First, notice that B^* is normalized. Since M^* is a prior-independent β-approximation, $M^*(F) \geq \frac{1}{\beta}\operatorname{OPT}_F(F)$ for all F in the family of distributions. Multiplying through by β and applying equation (2) shows that the benchmark meets the definition of normalization.

Second, equation (1) implies that M^* is a prior-free β-approximation of B^*. Thus, (M^*, B^*) is a solution to the benchmark design program (γ) with ratio β. The optimal solution to the program is no larger. Thus, $\gamma \leq \beta$. ■

The benchmark design problem presented above asks for both an optimal benchmark and the optimal mechanism for this benchmark. Fixing the benchmark, the problem of identifying the optimal mechanism is not well understood. There is a canonical method for identifying lower bounds on the approximation ratio of the optimal mechanism. Goldberg et al. (2006) suggest that the prior-free approximation any mechanism for a benchmark can be lower bounded by identifying a distribution over inputs for which all non-dominated mechanisms obtain the same performance. For revenue maximizing mechanism design, this distribution is the so-called *equal revenue* distribution. The following lemma and definition generalize this lower bound to environments where a mechanism neutralizing distribution may not exist.

Lemma 2. *For any benchmark B, distribution F, and family of mechanisms which induce OPT_F, the optimal prior-free approximation ρ^M is at least $\frac{B(F)}{\operatorname{OPT}_F(F)}$.*

Proof: Let M^B be the prior-free optimal mechanism for benchmark B, i.e., that optimizes the program (ρ^B), then

$$\rho^B = \max_{\boldsymbol{v}\in\mathcal{V}} \frac{B(\boldsymbol{v})}{M^B(\boldsymbol{v})} \geq \frac{B(F)}{M^B(F)} \geq \frac{B(F)}{\operatorname{OPT}_F(F)}.$$

■

Definition 6. *For any benchmark B, family of distributions $\mathcal{F}$, and family of mechanisms which induce OPT_F, the* canonical lower-bound on the optimal prior-free approximation of B *is*

$$\bar{\rho}^B = \max_{F\in\mathcal{F}} \frac{B(F)}{\operatorname{OPT}_F(F)}. \tag{$\bar{\rho}^B$}$$

This lower bound holds for any family of distributions; e.g., contrasting to standard assumptions in mechanism design, it allows distributions that are irregular and correlated. Moreover, the lower bound is tight for a number of interesting benchmarks. By Theorem 1, the optimal benchmark B^* satisfies $\rho^{B^*} = \bar{\rho}^{B^*}$. For the digital goods revenue maximization problem, a large family of benchmarks were shown by Chen et al. (2014) to also satisfy this equality. Thus, a natural relaxation of the benchmark design program is, instead of optimizing benchmarks that admit the best prior-free approximation, to optimize benchmarks that admit the best lower bound of Definition 6.

Definition 7. *The* relaxed benchmark design *problem for distribution family $\mathcal{F}$ and family of mechanisms that defines the Bayesian optimal mechanism OPT_F for and distribution $F \in \mathcal{F}$ solves the program:*

$$\bar{\gamma} = \min_{B\in\mathcal{B}(\mathcal{F})} \max_{F\in\mathcal{F}} \frac{B(F)}{\operatorname{OPT}_F(F)}. \tag{$\bar{\gamma}$}$$

Proposition 2. *The value of the relaxed program lower bounds optimal resolution, i.e., $\bar{\gamma} \leq \gamma$.*

We will see in Section IV that the relaxation is not generally without loss and, in particular, the optimal benchmark for the relaxed program can have $\bar{\gamma} < \gamma$. We will show the strictness of this inequality by example for the paradigmatic problem of maximizing revenue from the sale of an item to one of two agents and benchmarks that are normalized for value distributions that are i.i.d. and regular. We show this inequality is strict for online learning as well.

III. Prior-Independent Optimal Mechanisms

In this section we consider the problem of maximizing revenue from the sale of a single item to one of two agents with values distributed independently and identically from a distribution that satisfies a natural convexity property (to be defined formally). We identify the prior-independent optimal mechanism and thus, by the equivalence between benchmark design and prior-independent optimization, the optimal benchmark. The section begins with preliminary discussion of mechanism design.

A. Mechanism Design Preliminaries

Consider n agents with private values $\boldsymbol{v} = (v_1, \ldots, v_n)$. The agents have linear utility given, e.g, agent i's utility is $v_i\,x_i - p_i$ for allocation probability x_i and expected payment p_i. Agents' values are drawn independently and identically from a product distribution $\boldsymbol{F} = F \times \cdots \times F$ where F will denote the cumulative distribution function of each agent's value.

A mechanism M is defined by an ex post allocation and payment rule $\boldsymbol{x}^M$ and $\boldsymbol{p}^M$ which map the profile of values $\boldsymbol{v}$ to a profile of allocation probabilities and a profile of payments, respectively. We focus on mechanisms that are feasible, dominant strategy incentive compatible, and individually rational:

- For selling a single item, a mechanism is *feasible* if for all valuation profiles, the allocation probabilities sum to at most one, i.e., $\forall \boldsymbol{v},\ \sum_i x_i^M(\boldsymbol{v}) \leq 1$.
- A mechanism is *dominant strategy incentive compatible* if no agent i with value v_i prefers to misreport some value z: $\forall \boldsymbol{v}, i, z,\ v_i\,x_i^M(\boldsymbol{v}) - p_i^M(\boldsymbol{v}) \geq v_i\,x_i^M(z, \boldsymbol{v}_{-i}) - p_i^M(z, \boldsymbol{v}_{-i})$ where $(z, \boldsymbol{v}_{-i})$ denotes the valuation profile with v_i replaced with z.
- A mechanism is *individually rational* if truthful reporting always leads to non-negative utility: $\forall \boldsymbol{v}, i,\ v_i\,x_i^M(\boldsymbol{v}) - p_i^M(\boldsymbol{v}) \geq 0$.

A mechanism's revenue can be easily and geometrically understood via the marginal revenue approach of Myerson (1981) and Bulow and Roberts (1989). For distribution F, the *quantile* q of an agent with value value v denotes how strong that agent is relative to the distribution F. Quantiles are defined by the mapping $Q_F(v) = 1 - F(v)$. Denote the mapping back to value space by V_F, i.e., $V_F(q)$ is the value of the agent with quantile q. A single agent *price-posting revenue curve* gives the revenue of posting a price as a function of the probability that the agent accepts the price. For an agent with value distribution F, price $V_F(q) = F^{-1}(1 - q)$ is accepted with probability q; its revenue is $q\,V_F(q)$. A single agent *revenue curve* gives the optimal revenue from selling to a single agent $R_F(q)$ as a function of ex ante sale probability q. Note that the revenue curve R is always concave. The agent is *regular* if her revenue curve equals her price-posting revenue curve. The optimal mechanism for a single agent posts the *monopoly price* $V_F(\hat{q}^\star)$ which corresponds to the monopoly quantile $\hat{q}^\star = \operatorname{argmax}_q R_F(q)$. The expected revenue of a multi-agent mechanism M is equal to its surplus of marginal revenue.

Theorem 2 (Myerson, 1981)**.** *Given any incentive-compatible mechanism M with allocation rule $\boldsymbol{x}^M(\boldsymbol{v})$, the expected revenue of mechanism M for agents with regular distribution $\boldsymbol{F}$ is equal to its expected surplus of marginal revenue, i.e.,*

$$\begin{aligned} M(\boldsymbol{F}) &= \sum\nolimits_i \mathbf{E}_{\boldsymbol{v}\sim \boldsymbol{F}}\big[p_i^M(\boldsymbol{v})\big] \\ &= \sum\nolimits_i \mathbf{E}_{\boldsymbol{v}\sim \boldsymbol{F}}\big[R_F'(Q_F(v_i))\,x_i^M(\boldsymbol{v})\big]\,. \end{aligned}$$

Corollary 2 (Myerson, 1981)**.** *For i.i.d., regular, single-item environments, the optimal mechanism OPT_F is the second-price auction with reserve equal to the monopoly price.*

Proof: Optimizing marginal revenue pointwise the item is assigned to the agent with the highest non-negative marginal revenue. Since agents are i.i.d. and the marginal revenue curves are monotonically non-increasing, this winning agent is the one with the highest value that exceeds the monopoly price. ■

The following lemma from Dhangwatnotai et al. (2015) follows from Theorem 2 and gives a geometric understanding of revenue in two-agent auctions.

Lemma 3 (Dhangwatnotai et al., 2015)**.** *In i.i.d. two-agent single-item environments, the expected revenue of the second price auction is twice the area under the revenue curve and the expected revenue of the optimal mechanism is twice the area under the smallest monotone concave upper bound of the revenue curve.*

B. Prior-independent Optimization

In the remainder of this section we solve for the prior-independent optimal mechanism for the revenue objective with the restriction to

- single-item, two-agent environments, i.e., $n = 2$ (implicit);
- the family of i.i.d. regular value distribution $\mathcal{F}^{\mathrm{Reg}}$; and
- the family of feasible, incentive compatible, individually rational, and scale-invariant mechanisms $\mathcal{M}^{\mathrm{SI}}$.

The following discussion motivates these restrictions. The single-item two-agent environment is canonical for prior-independent revenue maximization. There do not exist good prior-independent mechanisms for general asymmetric and irregularly distributed agent values. Almost all papers on prior-independent mechanism design restrict to i.i.d. agents. Almost all papers on revenue maximization for prior-independent mechanism design restrict to regular distributions. The restriction to feasible and individually rational mechanisms is required to have a sensible optimization problem. The restriction to incentive compatible mechanisms is made in almost all papers on prior-independent mechanism design, an exception is Feng and Hartline (2018) where it is shown that the restriction can be lossy. The remaining condition which we formally define below is scale invariance.

Definition 8. *Given any incentive-compatible mechanism M with allocation rule $x^M(\boldsymbol{v})$, mechanism M is* scale-invariant *if for each agent i, valuation profile $\boldsymbol{v}$ and any constant $\alpha > 0$, $x_i^M(\alpha \cdot \boldsymbol{v}) = x_i^M(\boldsymbol{v})$. Scale invariance further implies $M(a \cdot \boldsymbol{v}) = a \cdot M(\boldsymbol{v})$.*

Allouah and Besbes (2018) prove that the optimal prior-independent mechanism among a broad family of mechanisms is scale invariant. They show that if $\lim_{\alpha \to 0} x_i(\alpha \cdot \boldsymbol{v})$ always exists for mechanisms in the family, then the optimal prior-independent mechanism is scale invariant. They conjecture that this weaker assumption is without loss; if true, the mechanism we identify as the optimal mechanism among scale-invariant mechanisms is also prior-independent optimal among all mechanisms.

Given the restriction to scale-invariant mechanisms, it will be sufficient to consider distributions that are normalized so that the single-agent optimal revenue is $\max_q R(q) = 1$.

The following family of (stochastic) markup mechanisms is (essentially, in $n = 2$ agent environments) the restriction of the family of lookahead mechanisms (Ronen, 2001) to those that are scale invariant. Notice that the second-price auction is the 1-markup mechanism M_1.

Definition 9. *The r-*markup mechanism *M_r identifies the agent with the highest-value (and ties broken uniformly at random) and offers this agent r times the second-highest value. A* stochastic markup mechanism *draws r from a given distribution on $[1, \infty)$. The family of stochastic markup mechanisms is $\mathcal{M}^{\mathrm{SMKUP}}$.*

Theorem 3. *For i.i.d., regular, two-agent, single-item environments, the optimal scale-invariant, incentive-compatible mechanism for prior-independent optimization program (β) is M_{α^*,r^*} which randomizes over the second-price auction M_1 with probability α^* and r^*-markup mechanism M_{r^*} with probability $1 - \alpha^*$, where $\alpha^* \approx 0.806$ and $r^* \approx 2.447$. The worst-case regular distribution for this mechanism is triangle distribution $\mathrm{Tri}_{\bar{q}^*}$ with $\bar{q}^* \approx 0.093$ and its approximation ratio is $\beta \approx 1.907$.*

In the two sections below we prove this theorem with the following main steps. First, we characterize the prior-independent optimal mechanism under the restriction to stochastic markup mechanisms and triangle distributions, cf. Alaei et al. (2018). This restricted program has the same solution as is given in Theorem 3. Second we show that the stochastic markup mechanisms and triangle distributions are mutual best responses among the more general families of scale-invariant mechanisms and regular distributions. Combining these results gives the theorem.

C. Stochastic Markup Mechanisms versus Triangle Distributions

In this section we characterize the solution to the prior-independent optimization program restricted to stochastic markup mechanisms and triangle distributions. We first define triangle distributions, which have revenue curves shaped like triangles (Figure 1), as well as a more general family of truncated distributions, which will be important subsequently in the proof. Recall that for scale-invariant mechanisms, it is without loss to normalize the distributions to have monopoly revenue one.

Definition 10. *A* normalized triangle distribution with monopoly quantile *$\bar{q}$, denoted $\mathrm{Tri}_{\bar{q}}$, is defined by the quantile function*

$$Q_{\mathrm{Tri}_{\bar{q}}}(v) = \begin{cases} \frac{1}{1+v(1-\bar{q})} & v \leq 1/\bar{q} \\ 0 & \textit{otherwise.} \end{cases}$$

The triangulation *of a normalized distribution with monopoly quantile $\bar{q}$ is $\mathrm{Tri}_{\bar{q}}$. The family of normalized triangle distributions is $\mathcal{F}^{\mathrm{Tri}} = \{\mathrm{Tri}_{\bar{q}} : \bar{q} \in [0, 1]\}$.*

Definition 11. *A distribution is* truncated *if the highest-point in its support is the monopoly price (typically a point mass). The* truncation *of a distribution is the distribution that replaces every point above the monopoly price with the monopoly price. The family of truncated distributions is denoted $\mathcal{F}^{\mathrm{Trunc}}$.*

The three lemmas below give formulae for the revenue of the optimal mechanism, the second-price auction, and non-trivial markup mechanisms for triangle distributions. The formula for revenue of markup mechanisms is discontinuous at $r = 1$. Thus, in our discussion we will distinguish between the second-price auction M_1 and non-trivial markup mechanism M_r for $r > 1$.

Lemma 4. *For i.i.d., normalized truncated, two-agent, single-item environments, the optimal mechanism posts the monopoly price and obtains revenue $2 - \bar{q}$ where $\bar{q}$ is the probability that an agent's value equals the monopoly price.*

Proof: The smallest monotone concave function that upper bounds the revenue curve is a trapezoid; its area is $\bar{q}/2 + 1 - \bar{q}$. The optimal revenue from two agents, by Lemma 3, is twice this area, i.e., $2 - \bar{q}$. ■

Lemma 5. *The revenue of the second-price auction M_1 for distribution $\mathrm{Tri}_{\bar{q}}$ is 1, i.e., $M_1(\mathrm{Tri}_{\bar{q}}) = 1$.*

Proof: By Lemma 3, the revenue is twice the area under the revenue curve. That area is $1/2$; thus, the revenue is 1. ■

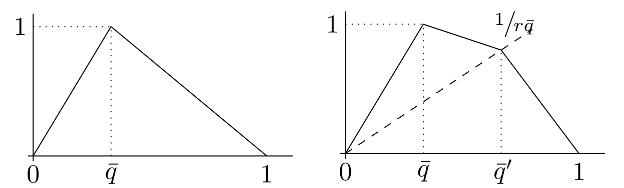

Figure 1. The left hand side is the revenue curve for triangle distribution $\mathrm{Tri}_{\bar{q}}$ and the right hand side is the revenue curve for quadrilateral distribution $\mathrm{Qr}_{\bar{q},\bar{q}',r}$. The definition of quadrilateral distribution $\mathrm{Qr}_{\bar{q},\bar{q}',r}$ will be formally introduced later in Section III-D.

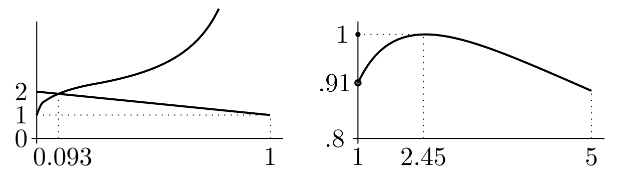

Figure 2. The figure on the left plots, as a function of $\bar{q}$, the approximation ratio $\mathrm{APX}_1(\bar{q})$ of the second-price auction M_1 against triangle distribution $\mathrm{Tri}_{\bar{q}}$ (straight line), and the approximation ratio $\mathrm{APX}_*(\bar{q})$ of the optimal non-trivial markup mechanism against triangle distribution $\mathrm{Tri}_{\bar{q}}$ (curved line). These functions cross at $\bar{q}^* = 0.0931057$. The figure on the right plots the revenue of the r markup mechanism M_r on triangle distribution $\mathrm{Tri}_{\bar{q}^*}$ as a function of markup r, i.e., $M_r(\mathrm{Tri}_{\bar{q}^*})$. Notice that, by choice of $\bar{q}^*$, the optimal non-trivial markup mechanism has the same revenue as the second-price auction.

Lemma 6. *The revenue of the r-markup mechanisms M_r on triangle distribution $\mathrm{Tri}_{\bar{q}}$, for $r \in (1, \infty)$ and $\bar{q} \in [0, 1)$, is*

$$M_r(\mathrm{Tri}_{\bar{q}}) = \frac{2r}{(1-\bar{q})(r-1)} \left(\frac{1-\bar{q}}{1-\bar{q}+\bar{q}r} + \frac{\ln\left(\frac{r}{1-\bar{q}+\bar{q}r}\right)}{1-r} \right).$$

The following theorem characterizes the prior-independent optimal stochastic markup mechanism against triangle distributions. The parameters of this optimal mechanism are the solution to an algebraic expression (cf. Lemma 6) that we are unable to solve analytically. Our proof will instead combine numeric calculations of select points in parameter space with theoretical analysis to rule out most of the parameter space. We can show that the expression is well-behaved and, thus, numeric calculation can identify near optimal parameters. Due to the lack of space, the discussion of this hybrid numerical and theoretical analysis can be found in the full version.

Before giving Theorem 4, we give context for its proof. In abstract terms, the prior-independent optimization program (β) can be viewed as a zero sum game between the designer and an adversary, where the designer chooses a prior-independent mechanism M, the adversary chooses a worst-case distribution F (and its induced revenue curve), and the payoff of the designer is the approximation ratio $\mathrm{OPT}_F(F)/M(F)$ (see Definition 2).

Theorem 4. *For i.i.d., triangle distribution, two-agent, single-item environments, the optimal stochastic markup mechanism for prior-independent optimization program (β) is M_{α^*,r^*} which randomizes over the second-price auction M_1 with probability α^* and r^*-markup mechanism M_{r^*} with probability $1-\alpha^*$, where $\alpha^* \approx 0.806$ and $r^* \approx 2.447$. The worst-case distribution for this mechanism is the triangle distribution $\mathrm{Tri}_{\bar{q}^*}$ with $\bar{q}^* \approx 0.093$ and its approximation ratio is $\beta \approx 1.907$.*

D. Mutual best-response of Stochastic Markup Mechanisms and Triangle Distributions

In this section we show that stochastic markup mechanisms are a best response (for the designer) to truncated distributions and that truncated distributions are a best response (for the adversary) to stochastic markup mechanisms. Moreover, we show that among truncated distributions, triangle distributions are the best for the adversary. Triangle distributions are known to be worst case for other questions of interest in mechanism design, e.g., approximation by anonymous reserves and anonymous pricings (Alaei et al., 2018). The proof that triangle distributions are worst-case for two-agent prior-independent revenue maximization is significantly more involved than these previous results.

Theorem 5. *For i.i.d., two-agent, single-item environments and any scale-invariant incentive-compatible mechanism M, there is a stochastic markup mechanism M' with (weakly) higher revenue (and weakly lower approximation ratio) on every truncated distribution F. I.e., $M'(F) \geq M(F)$.*

Proof: In a stochastic markup mechanism the price of the higher agent is a stochastic multiplicative factor $r \geq 1$ of the value of the lower agent (with ties broken randomly). To prove this theorem we must argue that (a) if the agents are not tied, then revenue improves if the lower agent loses, (b) if the agents are tied, then revenue is unaffected by random tie-breaking, and (c) any such scale-invariant mechanism looks to the higher-valued agent like a stochastic posted pricing with price that is a multiplicative factor (at least one) of the lower-valued agent's value. The detailed reason why arguments (a), (b) and (c) holds can be found in the full version of the paper. ■

Next we will give a sequence of results that culminate in the observation that for any regular distribution and any stochastic markup mechanism with probability α at least $2/3$ on the second-price auction (which includes the optimal mechanism from Theorem 4) either the triangulation of the distribution or the point mass Tri_1 has (weakly) higher approximation ratio. As the notation indicates, the point mass distribution Tri_1 is a triangle distribution.

Theorem 6. *For i.i.d., two-agent, single-item environments and any regular distribution F and any stochastic markup mechanism M that places probability $\alpha \in [2/3, 1]$ on the second-price auction, either the triangulation of the distribution F^{Tri} or the point mass Tri_1 has (weakly) higher approx-*

imation ratio. I.e., $\max\left\{\frac{\mathrm{OPT}_{F^{\mathrm{Tri}}}(F^{\mathrm{Tri}})}{M(F^{\mathrm{Tri}})}, \frac{\mathrm{OPT}_{\mathrm{Tri}_1}(\mathrm{Tri}_1)}{M(\mathrm{Tri}_1)}\right\} \geq \frac{\mathrm{OPT}_F(F)}{M(F)}$.

To prove this theorem we give a sequence of results showing that for any regular distribution, a corresponding truncated distribution is only worse; for any truncated distribution and a fixed stochastic markup mechanism (that mixes over M_1 and some M_r), a corresponding quadrilateral distribution (based on r) is only worse; and for any quadrilateral distribution, a corresponding triangle distribution (independent of r) is only worse. The theorem follows from combining these results. The first step assumes that the probability that the stochastic markup mechanism places on the second price auction is $\alpha \in [1/2, 1]$; the last step further assumes that $\alpha \in [2/3, 1]$.

To begin, the following lemma shows that the best response of the adversary to a relevant stochastic markup mechanism is a truncated distribution. Recall that by Fu et al. (2015) the prior-independent optimal mechanism is strictly better than a 2-approximation. On the other hand, any stochastic markup mechanism that places probability α on the second-price auction M_1 has prior-independent approximation at least $1/\alpha$. Specifically, on the (degenerate) distribution that places all probability mass on 1, a.k.a. Tri_1, the approximation factor of such a stochastic markup mechanism is exactly $1/\alpha$. We conclude that all relevant stochastic markup mechanisms place probability $\alpha > 1/2$ on the second-price auction. Thus, this lemma applies to all relevant mechanisms. Through the rest of this section, omitted proofs of lemmas are available in the full version.

Lemma 7. *For i.i.d., two-agent, single-item environments, any regular distribution F, and any stochastic markup mechanism M that places probability $\alpha \in [1/2, 1]$ on the second-price auction; either the truncation of the distribution F' or the point mass distribution Tri_1 has (weakly) higher approximation ratio. I.e.,* $\max\left\{\frac{\mathrm{OPT}_{F'}(F')}{M(F')}, \frac{\mathrm{OPT}_{\mathrm{Tri}_1}(\mathrm{Tri}_1)}{M(\mathrm{Tri}_1)}\right\} \geq \frac{\mathrm{OPT}_F(F)}{M(F)}$.

The revenue of the optimal mechanism and the second-price auction are decomposed into two parts by whether the quantile for the second highest price is above or below $\bar{q}$. This decomposition is illustrated in Figure 3. Then by simple algebra we can show that for any stochastic markup mechanism M that places probability $\alpha \in [1/2, 1]$ on the second-price auction, the approximation ratio is maximized when the distribution is truncated.

The next step is to show that, among truncated distributions, the worst-case distribution for stochastic markup mechanisms are those with quadrilateral-shaped revenue curves, i.e., ones that are piecewise linear with three pieces (see Figure 1). Recall that for a truncated distribution at monopoly quantile $\bar{q}$, the upper bound of the support is a point mass on $1/\bar{q}$.

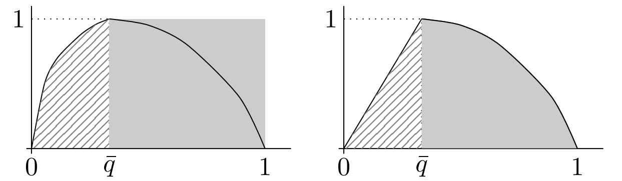

Figure 3. The illustration of the revenue decomposition of Lemma 7 for M on distribution F and truncation F' for the optimal mechanism and second-price auction. The thin black line on the left and right figures are the revenue curves corresponding to F and F', respectively. The dashed area on the left represents $\mathrm{OPT}_+ = \mathrm{SPA}_+$, the revenue from the optimal mechanism and second-price auction for distribution F when the quantile for the second highest price is above $\bar{q}$; and the gray area on the left represents $\mathrm{OPT}_- = \mathrm{OPT}'_-$, the revenue from the optimal mechanism for distribution F and F' when the quantile for the second highest price is below $\bar{q}$. The dashed area on the right represents $\mathrm{OPT}'_+ = \mathrm{SPA}'_+$, the revenue from the optimal mechanism and second-price auction for distribution F' when the quantile for the second highest price is above $\bar{q}$; and the gray area on the right represents $\mathrm{SPA}_- = \mathrm{SPA}'_-$, the revenue from the second-price auction for distribution F and F' when the quantile for the second highest price is below $\bar{q}$.

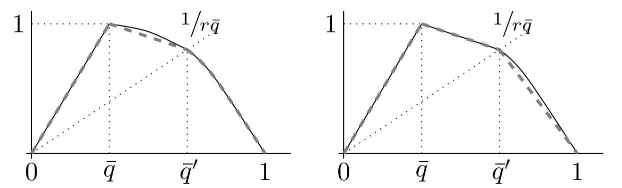

Figure 4. The main two steps of Lemma 9 are illustrated. In the first step (right-hand side), the revenue curves of distributions F^{Trunc} (thin, solid, black) and $F^{\dagger}$ (thick, dashed, gray) are depicted. In the second step, the revenue curves of the distributions $F^{\dagger}$ (thin, solid, black) and F^{Qr} (thick, dashed, gray) are depicted. In both cases the revenue of the r-markup mechanism is is higher on the thin, solid, black curve than the thick, dashed, gray curve.

Definition 12. *A* normalized quadrilateral distribution *with parameters $\bar{q}, \bar{q}'$ and r with $r \geq 1$ and $\frac{\bar{q}r}{\bar{q}r+(1-\bar{q})} \leq \bar{q}' \leq \min\{r\bar{q}, 1\}$, denoted by $\mathrm{Qr}_{\bar{q},\bar{q}',r}$ is defined by quantile function as:*

$$Q_{\mathrm{Qr}_{\bar{q},\bar{q}',r}}(v) = \begin{cases} \frac{\bar{q}'}{\bar{q}'+v r\bar{q}(1-\bar{q}')} & v < 1/r\bar{q} \\ \frac{\bar{q}'\bar{q}(r-1)}{v r\bar{q}(\bar{q}'-\bar{q})+(r\bar{q}-\bar{q}')} & 1/r\bar{q} \leq v \leq 1/\bar{q} \\ 0 & 1/\bar{q} < v \end{cases}$$

The following lemma summarizes an analysis from Allouah and Besbes (2018) and is useful in bounding the revenue from markup mechanisms.

Lemma 8 (Allouah and Besbes, 2018)**.** *Consider the r-markup mechanism, two i.i.d. regular agents with value distribution F, quantile $\bar{q}'$ corresponding to the monopoly price divided by r, and the distribution $\tilde{F}$ that corresponds to F ironed on $[\bar{q}', 1]$: the virtual surplus from quantiles $[\bar{q}', 1]$ is higher for F than for $\tilde{F}$.*

The next lemma reduces the worst case distribution from the family of truncated distributions to the family of quadrilateral distributions. The reduction is illustrated in Figure 4, by showing that ironing the revenue curves

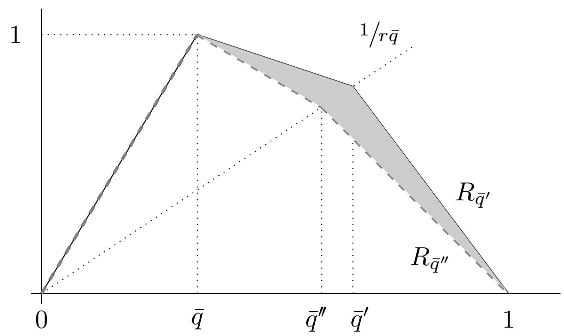

Figure 5. Illustrating the proof of Lemma 10, the difference of revenue for second price auction M_1 on revenue curves $R_{\bar{q}'}$ and $R_{\bar{q}''}$, which respectively correspond to quadrilateral distributions $\mathrm{Qr}_{\bar{q},\bar{q}',r}$ and $\mathrm{Qr}_{\bar{q},\bar{q}'',r}$, is equal to twice of the gray area, which is at least $\bar{q}'' - \bar{q}'$. Moreover, the difference of revenue for the r-markup mechanism M_r on revenue curves $R_{\bar{q}'}$ and $R_{\bar{q}''}$ is at most $2(\bar{q}'' - \bar{q}')$.

sequentially within $[\bar{q}, \bar{q}']$ and $[\bar{q}', 1]$ decreases the revenue of the stochastic markup mechanism. The optimal revenue is not affected because it is obtained using a reserve price corresponding to the monopoly quantile $\bar{q}$ and it is agnostic to the shape of the revenue curve for $q > \bar{q}$.

Lemma 9. *For i.i.d., two-agent, single-item environments, any truncated distribution F^{Trunc}, and any stochastic markup mechanism $M_{\alpha,r}$ with probability α on the second-price auction M_1 and probability $1-\alpha$ on non-trivial markup mechanism M_r; there is a quadrilateral distribution F^{Qr} with the same optimal revenue and (weakly) lower revenue in $M_{\alpha,r}$. I.e., $\mathrm{OPT}_{F^{\mathrm{Qr}}}(F^{\mathrm{Qr}}) = \mathrm{OPT}_{F^{\mathrm{Trunc}}}(F^{\mathrm{Trunc}})$ and $M_{\alpha,r}(F^{\mathrm{Qr}}) \leq M_{\alpha,r}(F^{\mathrm{Trunc}})$.*

We complete the proof of Theorem 6 by showing that triangle distributions lead to lower revenue than quadrilateral distributions. The intuition is illustrated in Figure 5. For any stochastic markup mechanism $M_{\alpha,r}$ with $\alpha \in [2/3, 1]$, consider a family of quadrilateral distributions $\mathrm{Qr}_{\bar{q},\bar{q}',r}$ parameterized by $\bar{q}'$. The optimal revenue is again not affected by $\bar{q}'$ while the revenue of $M_{\alpha,r}$ is monotone increasing in $\bar{q}'$. Thus the approximation ratio of $M_{\alpha,r}$ is maximized by minimal $\bar{q}'$ for which the degenerate quadrilateral is a triangle.

Lemma 10. *For i.i.d., two-agent, single-item environments, normalized quadrilateral distribution $\mathrm{Qr}_{\bar{q},\bar{q}',r}$, and stochastic markup mechanism $M_{\alpha,r}$ with probability $\alpha \in [2/3, 1]$ on the second-price auction M_1 and probability $1-\alpha$ on non-trivial markup mechanism M_r; the triangle distribution $\mathrm{Tri}_{\bar{q}}$ has the same optimal revenue and (weakly) lower revenue in $M_{\alpha,r}$. I.e., $\mathrm{OPT}_{\mathrm{Tri}_{\bar{q}}}(\mathrm{Tri}_{\bar{q}}) = \mathrm{OPT}_{\mathrm{Qr}_{\bar{q},\bar{q}',r}}(\mathrm{Qr}_{\bar{q},\bar{q}',r})$ and $M_{\alpha,r}(\mathrm{Tri}_{\bar{q}}) \leq M_{\alpha,r}(\mathrm{Qr}_{\bar{q},\bar{q}',r})$.*

IV. Sub-optimality of Relaxed Benchmark Design

In this section, we will show that the relaxed benchmark design program $(\bar{\gamma})$ is not generally equal to the benchmark design program (γ) by considering the revenue maximization problem for two agents with i.i.d. regular distributions. Since benchmark optimization and prior-independent mechanism design are equivalent problems (see Section III, Theorem 1), a gap between the objective values $\bar{\gamma}$ and γ of programs $(\bar{\gamma})$ and (γ) is implied by exhibiting a benchmark with lower objective value in program $(\bar{\gamma})$ than the approximation achieved by the optimal prior-independent mechanism, i.e., the solution to program (β).

Theorem 7. *For i.i.d., regular, two-agent, single-item environments and scale-invariant, incentive-compatible mechanisms the heuristic benchmark optimization program $(\bar{\gamma})$ has a strictly smaller objective value than the benchmark optimization program (γ), i.e., $\bar{\gamma} < \gamma$.*

The proof of Theorem 7 is deferred to the full version of the paper.

V. Prior-free versus Prior-independent Expert Learning

A main result of the paper, given in Section II, is that optimal benchmark design, as we have defined it, is equivalent to prior-independent optimization. Moreover, the optimal prior-free algorithm for the optimal benchmark is the optimal prior-independent algorithm. A consequence of these results is that there is no added robustness from the prior-free framework over the prior-independent framework. In this section we observe, by an example of expert learning, that this potential lack of robustness is serious and the optimal prior-independent algorithm can perform much worse than the standard algorithms that are known to approximate the standard prior-free benchmark. These observations are straightforward from the perspective of the expert learning literature; we discuss them in detail so as to map them onto the framework of Section II and give formal proofs for completeness. (Our full version shows that the relaxed benchmark design program is not without loss of generality.)

We consider the binary-reward variant of the canonical online expert learning problem. A single player plays a repeated game against Nature for n rounds. In each round t, each expert j from a discrete set $\{1, \ldots, k\}$ will receive a binary reward $v_{t,j} \in \{0, 1\}$. Thus, the input space is $\mathcal{V} = [\{0,1\}^k]^n$. Before rewards are realized, the player chooses to "follow" a (possibly randomized) expert for the round, and receives a reward (possibly in expectation) equal to the reward of the followed expert. When the round concludes, the player gets to observe the rewards of all experts, including those not followed by the player. The player's algorithm is M which outputs distributions $M_t(\boldsymbol{v})$ over experts using only the history $(v_1, \ldots, v_{t-1})$ in each round t. The class of all such *online* algorithms is denoted by $\mathcal{M}^{\mathrm{OL}}$ and the performance of an online algorithm $M \in \mathcal{M}^{\mathrm{OL}}$ on input $\boldsymbol{v}$ is:

$$M(\boldsymbol{v}) = \sum\nolimits_{t=1}^{n} \mathbf{E}_{j \sim M_t(\boldsymbol{v})}[v_{t,j}]\,.$$

As described in Section II we can define Bayesian, prior-independent, and prior-free versions of the expert learning problem. We summarize as follows:

- In the Bayesian model, the optimal algorithm is $\mathrm{OPT}_F = \mathrm{argmax}_{M \in \mathcal{M}^{\mathrm{OL}}} M(F)$.
  Consider the following family of *binary independent stationary* distributions $\mathcal{F}^{\mathrm{BIS}}$ for the Bayesian variant of the expert learning problem. For a distribution $F \in \mathcal{F}^{\mathrm{BIS}}$, each expert j's reward in each round is a Bernoulli random variable with mean f_j. The class $\mathcal{F}^{\mathrm{BIS}}$ is composed of all possible means $f_j \in [0,1]$. At each round t, the rewards are drawn independently from each other and from other rounds. Importantly the distribution of each expert's reward is identical across rounds. For binary independent stationary distributions $F \in \mathcal{F}^{\mathrm{BIS}}$, the optimal algorithm picks the expert with the highest ex ante probability $j^* = \mathrm{argmax}_j f_j$ and follows expert j^* in each round; its expected performance is

  $$\mathrm{OPT}_F(F) = n \max_j f_j.$$

- In the prior-free model, the optimal algorithm is the one that minimizes regret in worst-case over inputs $\boldsymbol{v} \in \mathcal{V}$ against a given benchmark B defined as

  $$\rho^B = \min_{M \in \mathcal{M}^{\mathrm{OL}}} \max_{\boldsymbol{v} \in \mathcal{V}} [B(\boldsymbol{v}) - M(\boldsymbol{v})].$$

  The *best-in-hindsight* benchmark for any reward profile $\boldsymbol{v} \in \mathcal{V}$ is

  $$B^{\mathrm{BIH}}(\boldsymbol{v}) = \max\nolimits_{j=1}^{k} \sum\nolimits_{t=1}^{n} v_{t,j}.$$

  A typical online analysis measures performance in terms of worst-case regret with respect to the best-in-hindsight benchmark B^{BIH}.
- In the prior-independent model, the optimal algorithm is the one that minimizes regret in worst-case over distributions $F \in \mathcal{F}$ against the optimal algorithm for the distribution

  $$\beta = \min_{M \in \mathcal{M}^{\mathrm{OL}}} \max_{F \in \mathcal{F}} [\mathrm{OPT}_F(F) - M(F)].$$

  We will be considering this question for binary independent stationary distributions $\mathcal{F}^{\mathrm{BIS}}$ where $\mathrm{OPT}_F(F)$ is as described above.

We observe next that the best-in-hindsight benchmark is normalized. (In fact, it is analogous to the normalized benchmark described by Hartline and Roughgarden (2008) for evaluating prior-free mechanisms.) Thus, an algorithm that is a prior-free approximation of the benchmark is also a prior-independent approximation algorithm (with the same bound on regret, cf. Proposition 1).

Lemma 11. *For inputs from binary independent stationary distributions $\mathcal{F}^{\mathrm{BIS}}$, the best-in-hindsight benchmark B^{BIH} is normalized, i.e., $B^{\mathrm{BIH}}(F) \geq \mathrm{OPT}_F(F)$, $\forall F \in \mathcal{F}^{\mathrm{BIS}}$.*

Proof: Given $F \in \mathcal{F}^{\mathrm{BIS}}$, the Bayesian optimal algorithm OPT_F selects the same expert in each round. The best-in-hindsight benchmark selects the single expert that is best for the realized input $\boldsymbol{v}$. Thus, for all $\boldsymbol{v} \in \mathcal{V}$, $B^{\mathrm{BIH}}(\boldsymbol{v}) \geq \mathrm{OPT}_F(\boldsymbol{v})$. Taking expectations we have the lemma. ■

We now show that the natural *follow-the-leader* algorithm, which in round t chooses a uniform random expert from the set of experts with highest total reward from the first $t-1$ rounds, is the prior-independent optimal algorithm.

Definition 13. *The* follow-the-leader *algorithm selects an expert uniformly at random from the set of experts with highest total reward from previous rounds:*

$$M_t^{\mathrm{FTL}}(\boldsymbol{v}) = U[\mathrm{argmax}_j \sum\nolimits_{t'<t} v_{t',j}].$$

Theorem 8. *For binary independent stationary distributions, the follow-the-leader algorithm is the prior-independent optimal online learning algorithm.*

Proof: Consider the Bayesian optimal online algorithm for the *uniform permutation prior* defined by probabilities $\{f^j\}_{j=1}^n$ that are assigned to experts via a uniform random permutation σ (i.e., the reward of expert j is Bernoulli with mean $f_j = f^{\sigma(j)}$). The theorem follows from the optimality of follow-the-leader for any uniform permutation prior.

We first argue that the follow-the-leader algorithm is optimal for the any uniform permutation prior. The optimal algorithm for the uniform permutation prior forms a posterior from the reward history at any time t and chooses the expert with the highest expectation under this posterior. Naturally, the experts with the highest expected reward under the posterior are the ones with the highest historical reward (proof of this claim is shown in the appendix of the full version). In other words, follow-the-leader is the Bayesian optimal algorithm for the uniform permutation prior.

As before denote by β the prior-independent optimal regret. To complete the proof, consider the probabilities $\{f^j\}_{j=1}^n$ for which the Bayesian optimal algorithm for the uniform permutation prior obtains the largest regret. Observe that the regret of the Bayesian optimal algorithm for this uniform permutation prior lower bounds β. On the other hand, the prior-independent regret of the follow-the-leader algorithm upper bounds β. The follow-the-leader algorithm obtains the same regret on all permutations and this regret equals the Bayesian optimal regret for the uniform permutation prior; i.e., the upper bound and the lower bound are equal. ■

As we have proved in Section II, benchmark optimization (γ) and prior-independent optimization (β) are the same. Thus, the optimal prior-free benchmark is the performance of the follow-the-leader algorithm scaled up by its prior-independent approximation factor. Moreover, the optimal mechanism for the optimal benchmark is the follow-the-leader algorithm itself. While we may have hoped for the prior-free analysis to lead to more robust algorithms than the prior-independent analysis, by optimizing benchmarks in the framework provided in Section II, we have lost all of this potential robustness. Specifically, the standard expert learning algorithms that have low worst-case regret against the best-in-hindsight benchmark exhibit robustness that the follow-the-leader lacks. This observation is formalized in the following lemma which contrasts with the optimal regret of standard algorithms like randomized weighted-majority (Littlestone and Warmuth, 1994). The optimal worst-case regret against best-in-hindsight is $\Theta(\sqrt{n \ln k})$ for k experts, n rounds, and binary rewards (Haussler et al., 1995).

Lemma 12. *The prior-free regret of follow-the-leader against the best-in-hindsight benchmark with n rounds is $\Theta(n)$.*

Proof: Consider the input with an even number of rounds and rounds alternating as:

- odd round payoffs: $(1, 0, \ldots, 0) \in \{0, 1\}^k$,
- even round payoffs: $(0, 1, \ldots, 1) \in \{0, 1\}^k$.

The follow-the-leader algorithm chooses a uniform random expert for odd rounds and obtains expected payoff $\frac{1}{k}$ and chooses expert 1 for even rounds and obtains expected payoff of 0. The total expected payoff of follow the leader is $M^{\text{FTL}}(\boldsymbol{v}) = \frac{n}{2k}$. On the other hand, the best-in-hindsight benchmark is $B^{\text{BIH}}(\boldsymbol{v}) = \frac{n}{2}$. The additive regret is $B^{\text{BIH}}(\boldsymbol{v}) - M^{\text{FTL}}(\boldsymbol{v}) = \frac{n}{2}(1 - \frac{1}{k}) \in \Theta(n)$. ■

The observations of this section suggest that further study of the formulation of the benchmark optimization problem is necessary to better understand the trade-offs between prior-free and prior-independent robustness.

ACKNOWLEDGMENT

Work done by all authors was supported in part by NSF CCF 1618502.

REFERENCES

Alaei, S., Hartline, J., Niazadeh, R., Pountourakis, E., and Yuan, Y. (2018). Optimal auctions vs. anonymous pricing. *Games and Economic Behavior*.

Allouah, A. and Besbes, O. (2018). Prior-independent optimal auctions. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 503–503. ACM.

Borodin, A. and El-Yaniv, R. (2005). *Online computation and competitive analysis*. cambridge university press.

Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *The Journal of Political Economy*, 97:1060–90.

Chen, N., Gravin, N., and Lu, P. (2014). Optimal competitive auctions. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 253–262, New York, NY, USA. ACM.

Devanur, N. R., Hartline, J. D., and Yan, Q. (2015). Envy freedom and prior-free mechanism design. *Journal of Economic Theory*, 156:103–143.

Dhangwatnotai, P., Roughgarden, T., and Yan, Q. (2015). Revenue maximization with a single sample. *Games and Economic Behavior*, 91:318–333.

Feng, Y. and Hartline, J. D. (2018). An end-to-end argument in mechanism design (prior-independent auctions for budgeted agents). In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 404–415. IEEE.

Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139.

Fu, H., Immorlica, N., Lucier, B., and Strack, P. (2015). Randomization beats second price as a prior-independent auction. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 323–323. ACM.

Goldberg, A. and Hartline, J. (2003). Envy-free auctions for digital goods. In *Proc. 4th ACM Conf. on Electronic Commerce*, pages 29–35. ACM Press.

Goldberg, A. V., Hartline, J. D., Karlin, A., Saks, M., and Wright, A. (2006). Competitive auctions. *Games and Economic Behavior*, 55:242–269.

Hartline, J. and McGrew, R. (2005). From optimal limited to unlimited supply auctions. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 175–182.

Hartline, J. and Roughgarden, T. (2008). Optimal mechanism design and money burning. In *Proc. 39th ACM Symp. on Theory of Computing*, pages 75–84.

Haussler, D., Kivinen, J., and Warmuth, M. K. (1995). Tight worst-case loss bounds for predicting with expert advice. In Vitányi, P., editor, *Computational Learning Theory*, pages 69–83, Berlin, Heidelberg. Springer Berlin Heidelberg.

Koutsoupias, E. and Papadimitriou, C. H. (2000). Beyond competitive analysis. *SIAM Journal on Computing*, 30(1):300–317.

Littlestone, N. and Warmuth, M. K. (1994). The weighted majority algorithm. *Information and computation*, 108(2):212–261.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6:58–73.

Ronen, A. (2001). On approximating optimal auctions. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 11–17. ACM.

Sleator, D. D. and Tarjan, R. E. (1985). Amortized efficiency of list update and paging rules. *Communications of the ACM*, 28(2):202–208.