Fairness in limited resource prediction-driven decisions

Eitan Bachmat
Faculty of computer and information science
Ben-Gurion University of the Negev
ebachmat@bgu.ac.il

Inbal Livni Navon
Faculty of computer and information science
Ben-Gurion University of the Negev
inballn@bgu.ac.il

Abstract

In recent years, many high-stakes societal decisions are made by machine learning systems, often under strict capacity constraints that limit resources to a small subset of individuals. In such settings, information gaps across populations can lead to highly unbalanced allocations. We study prediction-driven resource allocation through the lens of fairness. We connect classical algorithmic fairness notions with resource allocation definitions, and characterize the trade-offs between fairness and utility. We introduce an adaptation of proportional fairness to this setting, showing that it yields a continuum of fairness criteria via a regularization parameter, along with quantitative bounds on the resulting price of fairness. We further show that max-min fairness and equal opportunity can incur an unbounded price of fairness in extreme cases, and propose a variant of equal opportunity with a bounded price of fairness.

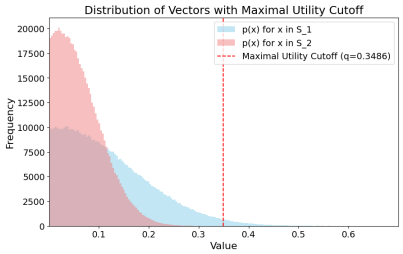
1 Introduction

In recent years, machine-learning algorithms have been used in many applications across various settings. Many important use cases arise in limited-resource settings, where predictions can be used to prioritize individuals for interventions or resources [YNT21, CB21, ANW20]. Such capacity constraints arise naturally in many domains. In medicine, predictions can help prioritize patients for limited healthcare resources. In education, predictions can be used to identify students at risk who should be prioritized for additional support [CB21], while special programs for high-potential students often also have strict capacity. Similarly, institutions allocating funding, such as research grants, scholarships, or start-up investments, have fixed budgets and must select among competing applicants or projects.

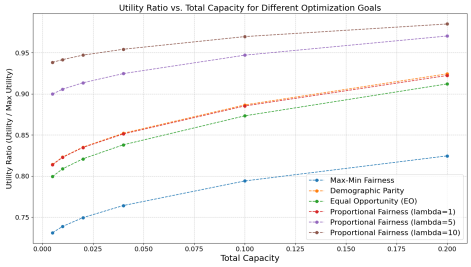
In all of these settings, the machine-learning decision involves two different considerations - risk prediction and resource allocation. The two considerations exist both when the machine-learning algorithm returns the allocation itself and when the algorithm is a risk predictor that we threshold to obtain an allocation. In the latter case, the allocation and risk prediction are clearly separated, but they both co-exist implicitly when the machine-learning algorithm returns the allocation itself. When discussing the quality of the decision, we aim to discuss both options, where in the case of direct allocation functions, the prediction is implicit.

When the resource is very limited and the risk prediction has large uncertainty, gaps in the quality of the prediction across different populations can cause a large discrepancy in the allocation of resources across these groups. In the absence of fairness considerations, if we can better recognize the highest-risk individuals from a group S_1, the algorithm may allocate almost all of the resources to individuals that belong to S_1, see Figure 1a. This can occur, for example, as a result of information gaps between different groups. Such situations are often unwanted by the decision makers. More generally, fairness considerations may be desirable whenever the individuals with the highest predicted risk or

40th Conference on Neural Information Processing Systems (NeurIPS 2026).



(a) Simulating the output distribution of a predictor p with a larger variance on S_1 (in blue) compared to S_2. The red line is a cutoff of 1% of the population, and nearly all individuals above the cutoff are from S_1. See Appendix F for more information.



(b) The ratio between the utility of the different fairness definitions to the maximal utility allocation, with different capacities. The underlying predictor distribution is as in Figure 1a. See Appendix F for details.

potential belong predominantly to a single group, regardless of whether this concentration results from information gaps or other differences between the groups. If a city develops a special program for gifted children, we do not want all of the children to come from the same school or neighborhood.

We consider settings in which the highest-risk or highest-potential individuals are disproportionately concentrated in one group, a property we formalize as tail dominance, and in which decision makers have decided to promote group fairness or diversity in the allocation. Our goal is to study how such fairness considerations affect the allocation of limited resources and the resulting utility. Importantly, our analysis does not depend on the underlying reason for the tail dominance; information gaps are one important mechanism through which it can arise.

Information gaps between individuals belonging to different groups appear in multiple important settings. In the medical setting, they can result from differences in the quality of medical history between individuals from different demographic groups [ZOP24]. In education, they may result from the structure of the school system or students' behavior [BCS25]. In sports, we have more information on the potential of older children compared to younger ones [BGMH24]. When allocating grants in research, newly appointed faculty members have larger uncertainty compared to older and more established ones. When decisions are made by a professional authority, they sometimes compensate for the lack of information. If we are to consider algorithmic decisions, we should formalize our goals in compensating information gaps.

In the algorithmic fairness literature, there is a large body of work discussing bias in algorithms and proposing different methods for mitigating it. One specific approach that is aimed at compensating for information gaps between populations is the equal opportunity fairness requirement [HPS16]. This requirement was originally developed for the general setting, without a strict capacity. In our work, we show that it has limitations when applied to our setting of strict capacity constraints.

1.1 Our contributions

Our goal in this work is to develop a theoretical understanding of group fairness in the allocation of limited resources based on individual risk.

Adapting resource allocation definitions In Section 3 we adapt fairness definitions from the resource-allocation literature to the setting of allocating resources to individuals based on their predicted risk. In classical resource allocation, a limited resource is divided among agents $i \in [m]$, each associated with a utility function u_i [BFT11, BOT18]. In our setting, the allocation is made at the individual level, while fairness is imposed across protected groups. We bridge these settings by identifying each protected group with an agent and defining its utility under an allocation f as the expected number of positive-outcome individuals receiving the resource within the group, $u_i = \mathbb{E}[f(x)y \mid x \in S_i]$. This choice captures the utility relevant to applications such as testing patients who end up positive or admitting students who will succeed. We then adapt max-min and proportional fairness to these group utilities. For proportional fairness, we additionally modify the normalization to account for the fact that each "agent" represents a population rather than a single individual.

Connecting group fairness definitions to resource allocation fairness In Section 4 we show that in a fully utilized setting, with equal base rates and rich enough hypothesis class, equal opportunity and our adapted max-min fairness are equivalent. Thus, in this setting, a standard group-fairness constraint can be understood as optimizing a classical resource-allocation fairness objective.

For proportional fairness, we provide a characterization specific to our group-allocation setting: proportional fairness can be viewed as a form of multiplicative regularization balancing total utility and equality between group utilities. In particular, satisfying proportional fairness is equivalent to maximizing an expression consisting of the log of the total utility and a term measuring deviation from equal group utility through the Kullback–Leibler divergence. This characterization motivates a parameterized family of proportional-fairness objectives that allows the strength of the fairness preference to vary.

Bounding the price of fairness In Section 5, we analyze the utility cost of imposing these fairness objectives, with particular attention to the tail-dominance setting motivating our work. For proportional fairness, we derive an explicit quantitative upper bound on the price of fairness, depending on the group sizes and the regularization parameter. This provides a direct way to choose the strength of the fairness objective while guaranteeing a desired bound on its utility cost.

In contrast, we show that max-min fairness and, through our equivalence result, equal opportunity with equal base rates, can have an arbitrarily large price of fairness under tail dominance. Enforcing these notions can shift resources away from the group containing most of the highest-risk individuals toward the other group, substantially reducing the total utility. We complement this qualitative separation with quantitative lower bounds on the price of fairness. Together, these results distinguish fairness notions that provide controllable utility guarantees from notions whose cost can become arbitrarily large precisely in the limited-resource regimes motivating our setting.

We complement our theoretical results with simulations comparing the price of fairness under different underlying predictor distributions in Appendix F. Figure 1b illustrates how the price of fairness of the different fairness definitions varies with capacity, using the predictor distribution shown in Figure 1a.

Suggesting new fairness definitions In addition to adapting proportional fairness and suggesting a generalization of it using a regularization parameter, in Section 6 we also suggest an adaptation of equal opportunity to the limited resources setting, and show that it has a bounded price of fairness.

1.2 Related works

The problem of fair allocations of resources is a rich research area in many different settings, such as networks [OLP+14], facility locations [MS94], division of goods [ABFRV22], operations research [BFT11] and more. These very different settings still share many of the fairness definitions used, such as max-min fairness.

In this work, we focus on fairness in resource allocation, where the allocation is based on individual risk prediction. There are works focusing on fairness in allocation combining other types of learning settings, such as predicting demand [DK20], learning gradually revealed utilization [EJJ+19], federated learning [ZMCY22], networks [WWW+18] and additional areas.

In fairness in machine-learning, the goal is to ensure that the algorithm is not biased against any group or individual. Hardt et al. [HPS16] propose the definition of equal opportunity to compensate for information gaps between groups. For risk prediction, [HJKRR18], suggest extending calibration to many overlapping groups. Gopalan et al. [GKR+21] show that a multicalibrated predictor can be post-processed to minimize different losses in an optimal way, and Hu et al. [HNRY23] show that it can be done under capacity and fairness constraints. Based on this framework, in parts of this work we model the allocation process as finding a multicalibrated predictor and post-processing it.

There are different works focusing on the price of enforcing different fairness definitions. In resource allocation [BFT11], it is usually modeled by comparing the utility with and without fairness constraints. In machine-learning [MW18, CDPF+17], it can be modeled both by the harm to the accuracy, or by defining a utility function, as we do in this work.

Recent works also consider fairness notions that apply only to specific regions of the score or outcome distribution [LDH26, CDE+26, HHYL25]. These works are related to ours in restricting fairness

requirements to a relevant part of the distribution, including its tail. Our setting differs in that the relevant region is induced by limited capacity: only a small fraction of the population can receive the resource, making the upper tail of the risk distribution particularly consequential. Accordingly, our focus is on fairness in capacity-constrained allocation and on quantifying the resulting fairness–utility trade-offs through the price of fairness.

When we use resource allocation fairness definitions such as max-min and proportional fairness, we model the utility as the true positive count. For a given capacity and base rates, maximizing the true positive count is equal to minimizing the loss, although our definitions of max-min and proportional fairness are not equivalent to loss equality. Modeling the utility as the loss is done in the literature, for example in [DOBD$^+$18, DGK$^+$21, ZMCY22].

1.3 Limitations and broader impact

Our work studies fairness in capacity-constrained decision-making from a theoretical perspective, proposing new definitions and quantitative bounds. We aim to broaden the framework for analyzing fairness–utility trade-offs and encourage fairness considerations in resource allocation problems often driven solely by utility.

Several steps are needed before applying our work in practice, requiring collaboration with domain experts. First, one must identify settings where group fairness interventions are desirable, such as when information gaps in healthcare cause certain populations to be overlooked for additional testing, or when only students from certain populations are identified as high-potential. Once an intervention is deemed appropriate, our framework allows decision makers to explore fairness–utility trade-offs and select an allocation policy with an acceptable price of fairness. However, as with any formal fairness notion, our definitions can be misused, for example, by choosing weak constraints or tuning parameters (such as the regularization weight in proportional fairness) to prioritize utility while nominally satisfying fairness requirements. Careful domain-specific judgment is therefore essential in selecting and applying these definitions.

Our work focuses on the design and analysis of fairness criteria rather than their implementation, and thus does not address online decision-making, learning a predictor from finite sample, or discrete capacity constraints. Results relying on predictor post-processing also depend on the quality and calibration of the underlying model, with guarantees degrading as prediction error increases, as reflected in our bounds.

Finally, our analysis uses randomization to break ties between individuals with identical predicted risk, but this is not essential: other tie-breaking criteria or "softer" randomized thresholds around the cutoff could be used. More generally, limited resources inevitably require choices among individuals near the allocation cutoff, introducing some arbitrariness. Our results are not tied to random tie-breaking and apply to other mechanisms for resolving such choices.

These considerations underscore the importance of combining theoretical guarantees with domain-specific judgment in practical applications.

2 Preliminaries and notations

We study prediction-based resource allocation under limited capacity. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$, where $x \in \mathcal{X}$ represents individual features and $y \in \{0,1\}$ is an unknown binary outcome. The goal is to allocate resources to individuals with high likelihood of $y = 1$. Unless stated otherwise, all expectations and probability distributions over x, y are under $\mathcal{D}$. We use $x \sim \mathcal{D}$ for the marginal distribution of x under $\mathcal{D}$. We assume the population is partitioned into m disjoint groups $S_1, \ldots, S_m$, and let $g : \mathcal{X} \to [m]$ denote the group membership function.

We allow randomized allocations. Let $f : \mathcal{X} \to [0,1]$, where $f(x)$ denotes the probability of allocating the resource to individual x.

Definition 2.1. A function $f : \mathcal{X} \to [0,1]$ has capacity c under $\mathcal{D}$ if $\mathbb{E}_{x\sim\mathcal{D}}[f(x)] = c$. For a group $S_i \subset \mathcal{X}$, f has capacity c on S_i if $\mathbb{E}_{x\sim\mathcal{D}}[f(x) \mid x \in S_i] = c$.

Our focus is on group fairness under information gaps between groups. We model decision-making as a two-step process: first, we learn a predictor $p : \mathcal{X} \to [0,1]$ for y, and second, we select individuals

by post-processing p. By [HNRY23], a predictor that is multicalibrated with respect to a function class $\mathcal{F}$ can be post-processed to satisfy group fairness constraints with loss matching the best function in $\mathcal{F}$ (up to calibration error). Thus, this two-step approach does not inherently degrade performance. In this framework, we model information gaps via *tail dominance*: the predictor p identifies high-risk individuals more effectively in one group than another.

Definition 2.2. Let $p : \mathcal{X} \to [0,1]$ and let $S_1, S_2 \subset \mathcal{X}$ be disjoint groups. We say that S_1 has t_0-tail dominance over S_2 under $\mathcal{D}$ if for all $r \geq t_0$,

$$\Pr_{x\sim\mathcal{D}}[p(x) > r \mid x \in S_1] \;\geq\; \Pr_{x\sim\mathcal{D}}[p(x) > r \mid x \in S_2].$$

Allocating resources to high-risk individuals can be implemented by post-processing a predictor $p : \mathcal{X} \to [0,1]$ using a post-processing function $\tau : [0,1] \times [m] \to [0,1]$. The resulting allocation is given by $f(x) = \tau(p(x), g(x))$. The post-processing may be arbitrary, but in this work we focus on randomized group thresholds, which allocate resources to the highest-risk individuals in each group.

Definition 2.3. A randomized group threshold function is a mapping $\tau : [0,1] \times [m] \to [0,1]$ such that for each group $i \in [m]$ there exist $t_i, \gamma_i \in [0,1]$ satisfying $\tau(r,i) = \mathbf{1}(r > t_i) + \gamma_i \mathbf{1}(r = t_i)$.

Under limited capacity, we seek to fully utilize the available resources. The following shows that this is always achievable using randomized thresholds.

Claim 2.4. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$, where $\mathcal{X}$ is partitioned into groups $S_1, \ldots, S_m$. For any $c \in [0,1]$ and group i with $\Pr_{x\sim\mathcal{D}}[x \in S_i] > 0$, there exists a randomized group threshold τ such that $\mathbb{E}_{x\sim\mathcal{D}}[\tau(p(x), i) \mid x \in S_i] = c$.

The proof is deferred to Appendix B. Some of the proofs in the paper use quantile functions. Appendix A provides preliminaries about quantile functions.

2.1 Loss minimization

In many decision-making settings, we minimize expected loss. For binary decisions and outcomes, let $\ell : \{0,1\} \times \{0,1\} \to [0,1]$ be a loss function. For a randomized assignment $f : \mathcal{X} \to [0,1]$, its expected loss under $\mathcal{D}$ is defined as follows (which reduces to $\mathbb{E}[\ell(f(x), y)]$ when f is binary).

Definition 2.5. The expected loss of a randomized assignment $f : \mathcal{X} \to [0,1]$ under $\mathcal{D}$ is given by $\mathbb{E}_{(x,y)\sim\mathcal{D}}\left[f(x)\ell(1,y) + (1 - f(x))\ell(0,y)\right]$.

Under full utilization, minimizing any binary loss ℓ is equivalent to maximizing the true positive count, defined below. Note that this is not the true positive rate, as it is unnormalized.

Definition 2.6. The true positive count of $f : \mathcal{X} \to [0,1]$ under $\mathcal{D}$ is $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y]$.

Claim 2.7. Let $\mathcal{D}$ be a distribution with $\Pr_{(x,y)\sim\mathcal{D}}[y = 1] \in (0,1)$. For any capacity $c \in [0,1]$, function class $\mathcal{F} \subseteq \{h : \mathcal{X} \to [0,1]\}$, and binary loss ℓ with $\ell(0,0) = \ell(1,1) = 0$, the function $f \in \mathcal{F}$ of capacity c that maximizes the true positive count is the function that minimizes the expected loss.

The proof is deferred to Appendix B. While some results are stated for general utilities, when instantiating a specific utility we use the true positive count, which corresponds to the total number of successful allocations.

2.2 Group fairness definitions

We introduce standard group fairness notions. For simplicity, we present exact definitions; our results extend to approximate versions.

Definition 2.8 (Demographic parity)**.** A function $f : \mathcal{X} \to [0,1]$ satisfies demographic parity with respect to $\mathcal{D}$ and $S_1, \ldots, S_m$ if for all $i, j \in [m]$, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_i] = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_j]$.

Definition 2.9 (Equal opportunity [HPS16])**.** A function $f : \mathcal{X} \to [0,1]$ satisfies equal opportunity with respect to $\mathcal{D}$ and groups $S_1, \ldots, S_m$ if for all $i, j \in [m]$,

$$\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_i,\, y = 1] = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_j,\, y = 1].$$

For predictors, a common requirement is calibration within groups. This applies to predictors p, where $p(x)$ estimates $\Pr[y = 1|x]$, and differs from allocations f, where $f(x)$ is the probability of allocating a resource to x. We remark that p can be both discrete or continuous, or mixed.

Definition 2.10 (Group calibration)**.** A predictor $p : \mathcal{X} \to [0, 1]$ is α-calibrated on groups $S_1, \ldots, S_m$ with respect to $\mathcal{D}$ if for all $i \in [m]$,

$$\mathop{\mathbb{E}}_{x' \sim \mathcal{D}, x' \in S_i} \left[\left| \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}} [y | x \in S_i, p(x) = p(x')] - p(x') \right| \right] \leq \alpha.$$

When post-processing a predictor to satisfy a fairness requirement, there are two approaches. The first treats the predictor p as correctly modeling $\Pr[y = 1|x]$, and enforces fairness with respect to the induced (simulated) distribution defined below. The second enforces fairness empirically using additional samples. If the distribution of $p(x)$ within each group is known, the first approach requires no new samples. However, if p is not well-calibrated, it may introduce significant error. In this work we bound the price of fairness using both approaches in different cases.

Definition 2.11. Let $\mathcal{D}$ be a distribution over $(x, y) \in \mathcal{X} \times \{0, 1\}$ and let $p : \mathcal{X} \to [0, 1]$. The simulated distribution $\mathcal{D}_p$ is defined by sampling $x \sim \mathcal{D}$ (marginally) and then $y \sim \mathrm{Ber}(p(x))$.

We quantify the cost of enforcing fairness via additive and multiplicative price of fairness notions.

Definition 2.12. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0, 1\}$ and let $\mathcal{F}$ be a class of functions. Given a fairness requirement, let $\mathcal{F}' \subseteq \mathcal{F}$ denote the subset of functions satisfying it. The price of fairness is

$$\text{Multiplicative: } \frac{\max_{h \in \mathcal{F}} \mathbb{E}_{(x,y) \sim \mathcal{D}}[h(x)y]}{\max_{f \in \mathcal{F}'} \mathbb{E}_{(x,y) \sim \mathcal{D}}[f(x)y]} \quad \text{Additive: } \max_{h \in \mathcal{F}} \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}}[h(x)y] - \max_{f \in \mathcal{F}'} \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}}[f(x)y].$$

3 Resource allocation fairness definitions in the prediction setting

In this section, we adapt classical resource allocation fairness notions to the prediction setting. We note that while we adopt definitions from resource allocation, we do not follow an axiomatic approach to utility. In the standard formulation, agents $i \in [m]$ are associated with utility functions u_i that map allocated resources to utility. We model each protected group $S_i \subset \mathcal{X}$ as an agent, with utility defined as the expected true positive count:

$$u_i(f) = \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}}[f(x)y \mid x \in S_i].$$

This choice of utility is a modeling decision. While Claim 2.7 shows that minimizing a binary loss is equivalent to maximizing the overall true positive count, this equivalence does not hold across groups with different base rates or capacities. As a result, fairness definitions based on loss are not equivalent to those based on true positive count.

We adopt true positive count as our utility since it captures the total benefit obtained from allocating the limited resource. For example, when the limited resource is a diagnostic test, i.e. $f(x) = 1$ is testing the patient x and y indicates whether they have the condition, the true positive count $f(x)y$ corresponds to the number of patients with the condition who receive the test and can therefore be correctly diagnosed. Similarly, in admissions, when y indicates whether a candidate would succeed if admitted, it captures the number of successful candidates who receive one of the limited slots. Thus, maximizing the true positive count corresponds to allocating the limited resource where it is expected to produce the largest number of successful outcomes. We remark that using the true positive rate as the utility is also possible, and the results are similar when the base rates are close.

Resource allocation fairness notions are optimization-based: they select allocations that optimize a function of the group utilities. In our setting, this corresponds to choosing $f \in \mathcal{F}$ to optimize a utility objective that jointly encodes both fairness and performance, and thus simultaneously enforces a fairness requirement and minimizes loss.

We adapt max-min fairness by defining group utilities via true positive count.

Definition 3.1. For a function class $\mathcal{F}$, a function $f : \mathcal{X} \to [0, 1]$ satisfies max-min fairness with respect to $\mathcal{D}$ and groups $\mathcal{S} = \{S_1, \ldots, S_m\}$ if

$$f = \arg\max_{h \in \mathcal{F}} \min_{i \in [m]} \mathop{\mathbb{E}}_{(x,y) \sim \mathcal{D}}[h(x)y \mid x \in S_i].$$

We also consider an adaptation of proportional fairness. In the classical formulation, utilities are defined per agent; here, agents correspond to groups, which introduces a normalization issue due to differing group sizes. In particular, splitting a group randomly into subgroups should not affect the fairness objective. To ensure this invariance, we use a normalized version that weighs each group by $\Pr[x \in S_i]$. We show in Appendix B that this definition is invariant under group splitting.

Definition 3.2. For a function class $\mathcal{F}$, a function $f : \mathcal{X} \to [0, 1]$ satisfies proportional fairness with respect to $\mathcal{D}$ and $\mathcal{S} = \{S_1, \ldots, S_m\}$ if

$$f = \arg\max_{h \in \mathcal{F}} \mathbf{Prop}_{\mathcal{D},\mathcal{S}}(h) = \arg\max_{h \in \mathcal{F}} \sum_{i \in [m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y \mid x \in S_i].$$

4 Connections between fairness definitions

We study the relationship between group fairness notions from algorithmic fairness and the resource allocation adaptations of fairness definitions introduced above.

4.1 Equal opportunity and max-min fairness

We characterize when equal opportunity and max-min fairness coincide. We begin by identifying sufficient conditions under which max-min fairness implies equal true positive count across groups (i.e., equal utility). Such equivalences are known in specific settings (e.g., [DOBD$^+$18]); we introduce an additional setting where this holds.

The two following conditions imply that max-min fairness equalizes the utility (true positive count) among all groups. While these conditions suffice to show it, there are also other settings for which max-min fairness equalizes the utility. See Appendix C for more details.

The first condition is non-degeneracy: allocating additional resources to any group strictly increases its expected true positive count. This condition typically holds under limited resources, where we do not have enough capacity to capture all of the $y = 1$ in the different groups.

Definition 4.1. A distribution $\mathcal{D}$ over $\mathcal{X} \times \{0, 1\}$ is non-degenerate with respect to groups $S_1, \ldots, S_m$, capacity c, and function class $\mathcal{F}$ if for every $f \in \mathcal{F}$ with $\Pr_{(x,y)\sim\mathcal{D}}[f(x) = 1] \leq c$ and every $i \in [m]$, we have $\mathbb{E}_{(x,y)\sim\mathcal{D}}[y \mid f(x) = 0,\ x \in S_i] > 0$.

The second condition is richness of the function class. A sufficient condition is that $\mathcal{F}$ is closed under group-wise post-processing. Notably, the set of post-processing of a predictor satisfies this property. There are other sets of functions $\mathcal{F}$ that equalizes the utility without satisfying the second condition, for example the set of randomized thresholds of a predictor. See Appendix C for more details.

Definition 4.2. A function class $\mathcal{F}$ is closed under group post-processing if for any $f \in \mathcal{F}$ and $\tau : [0, 1] \times [m] \to [0, 1]$, the function h defined by $h(x) = \tau(f(x), g(x))$ is in $\mathcal{F}$, where $g : \mathcal{X} \to [m]$ is the group-membership function.

Claim 4.3. For $c < 0.5$, let $\mathcal{D}$ be $2c$-non-degenerate with respect to groups $S_1, \ldots, S_m$, and let $\mathcal{F}$ be closed under group post-processing. Then any $f \in \mathcal{F}$ that satisfies max-min fairness with capacity c equalizes group utilities, i.e., for all $i, j \in [m]$,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y \mid x \in S_i] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y \mid x \in S_j].$$

The proof is deferred to Appendix C. It shows that if f does not equalize utilities, one can construct a group post-processing $h \in \mathcal{F}$ with a strictly larger max-min objective, contradicting optimality.

When groups have equal base rates, i.e., $\mathbb{E}[y \mid x \in S_i]$ is the same for all i, equal opportunity also enforces equal true positive counts. In this case, maximizing true positive count under equal opportunity coincides with max-min fairness.

Lemma 4.4. *For $c < 0.5$, let $\mathcal{D}$ be a non-degenerate distribution with respect to groups $S_1, \ldots, S_m$, capacity $2c$, and a function class that equalizes the group utilities. Let $\mathcal{F}_{EO} \subseteq \mathcal{F}$ denote the functions satisfying equal opportunity. If $\mathbb{E}[y \mid x \in S_i] = \mathbb{E}[y]$ for all $i \in [m]$, then*

$$\arg \max_{\substack{h \in \mathcal{F} \\ \text{capacity } c}} \min_{i \in [m]} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y \mid x \in S_i] = \arg \max_{\substack{h \in \mathcal{F}_{EO} \\ \text{capacity } c}} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y].$$

The proof is deferred to Appendix C; it shows that the optimizers of both problems coincide.

Non-equal base rates When base rates differ across groups, equal opportunity and max-min fairness are no longer equivalent. In particular, if a group S_i has a higher base rate than S_j, i.e., $\mathbb{E}[y \mid x \in S_i] > \mathbb{E}[y \mid x \in S_j]$, equal opportunity allocates more resources to S_i. This leads to higher utility of S_i comparing to S_j, whereas max-min fairness (under non-degeneracy) equalizes these quantities across groups. This difference arises from the choice of utility. If instead the utility of group i were defined as $\mathbb{E}[h(x)y \mid y = 1]$, then equal opportunity and max-min fairness would coincide even with unequal base rates (assuming non-degeneracy). The difference between these definitions becomes more pronounced as the gap in base rates increases.

4.2 Proportional fairness and fairness regularization

We show that normalized proportional fairness (Definition 3.2) decomposes into two terms: a Kullback–Leibler divergence between two distributions over groups, and the logarithm of the total true positive count. The first term is maximized when group utilities are equalized (as in max-min fairness), while the second is maximized by allocations that maximize overall true positive count. Thus, proportional fairness can be viewed as a form of fairness regularization. Unlike standard fairness regularization, this formulation is multiplicative rather than additive: both terms appear inside a logarithm, resulting in a multiplicative trade-off between fairness and performance at the multiplicative level.

We formalize the above decomposition by defining two distributions over groups. The first captures group weights under $\mathcal{D}$, and the second captures the distribution of the utility induced by allocation f. This part is in fact more general, and works for our normalized definition of proportional fairness for every group utility function $u_i(f)$ for an allocation f, where the total utility satisfies $U(f) = \sum_{i\in[m]} \Pr_{x\sim\mathcal{D}}[x \in S_i]u_i(f)$.

Definition 4.5. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0, 1\}$ with groups $\mathcal{S} = (S_1, \ldots, S_m)$.

1. The distribution $Q_{\mathcal{S}}$ over $[m]$ is defined by $\Pr_{j\sim Q_{\mathcal{S}}}[j = i] = \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]$.
2. The utility distribution Q_f over $[m]$ is defined by $\Pr_{j\sim Q_f}[j = i] = \frac{\Pr_{x\sim\mathcal{D}}[x\in S_i]u_i(f)}{U(f)}$.

Lemma 4.6. *For any $\mathcal{D}$, group partition $\mathcal{S} = \{S_1, \ldots, S_m\}$, and $f : \mathcal{X} \to [0, 1]$,*

$$\mathbf{Prop}_{\mathcal{D},\mathcal{S}}(f) = -D_{KL}(Q_{\mathcal{S}} \,\|\, Q_f) + \log U(f).$$

The proof is deferred to Appendix C. We can see that if a function f has $u_i(f) = u_j(f)$ for all groups i, j, then $Q_{\mathcal{S}} = Q_f$ and the divergence part zeros out. Therefore, from Claim 4.3 we have the following corollary. That is, max-min fairness zeros the fairness term in proportional fairness.

Corollary 4.7. *Let $\mathcal{D}$ be a $2c$ non-degenerate distribution with respect to groups $S_1, \ldots, S_m$, and let $\mathcal{F}$ be a function class closed under group post-processing. For $c < 0.5$, any $f \in \mathcal{F}$ with capacity c that satisfies max-min fairness also satisfies $D_{KL}(Q_{\mathcal{S}} \,\|\, Q_f) = 0$.*

Lemma 4.6 and Corollary 4.7 imply that proportional fairness is a form of multiplicative fairness regularization. Therefore, we can use a regularization factor λ and have a range of possible fairness definitions, where $\lambda = 0$ implies max-min fairness, $\lambda = 1$ proportional fairness, and larger λ values correspond to more relaxed fairness requirements.

Definition 4.8. For a function class $\mathcal{F}$ and regularization parameter $\lambda \geq 1$, a function $f : \mathcal{X} \to [0, 1]$ satisfies λ-proportional fairness with respect to $\mathcal{D}$ and $\mathcal{S}$ if

$$f = \arg\max_{h\in\mathcal{F}} -D_{KL}(Q_{\mathcal{S}} \,\|\, Q_h) + \lambda \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y].$$

In the next section we quantify the price of fairness for λ proportional fairness, as a function of λ.

5 The price of fairness

In this section, we quantify the price of fairness across different notions. To capture differences between groups, we model decision-making as a two-step process. The first is learning a predictor and the second is post-processing it. We quantify the information gap by tail dominance (Definition 2.2).

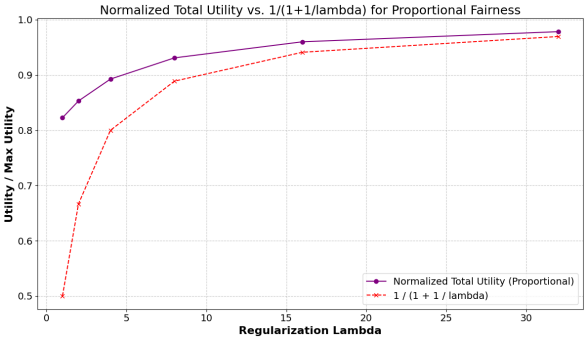


Figure 2: Simulation for the ratio between the utility of a λ-proportionally fair solution and the maximal utility solution, for various values of λ. In red is the lower bound of Theorem 5.1 on these parameters. See Appendix F for more information.

5.1 Proportional fairness

We show that proportional fairness has a bounded price of fairness. We assume that there is a majority group S_1 on whom we have the most information, and bound the price of fairness as a function of the size of the majority group and the regularization parameter.

Theorem 5.1. *Let $\mathcal{D}$ be a $2c$ non-degenerate distribution over $\mathcal{X} \times \{0,1\}$ where $\mathcal{X}$ is divided into disjoint groups $S_1 \ldots, S_m$. Let $p : \mathcal{X} \to [0,1]$ be an α-calibrated predictor and $\tau : [0,1] \times [m] \to [0,1]$ to be a randomized group threshold function such that $\tau(p(x), g(x))$ satisfies λ-proportional fairness with $\lambda \geq 1$ on the simulated distribution and has capacity c. Let τ' be the post-processing function such that $\tau'(p(x), g(x))$ maximizes the true positive count with capacity c. Suppose S_1 has t_0-dominance over all $S_i, i > 1$ for t_0 such that $\Pr[p(x) \geq t_0 \wedge x \in S_i] \geq c$ and that $\Pr[x \in S_1] \geq \mu$, then*

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x), g(x))y] \leq \left(1 + \frac{1-\mu}{\mu\lambda}\right) \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), g(x))y] + 4\alpha.$$

This theorem provides a quantitative bound on the price of fairness as a function of the regularization parameter and the size of the majority group. As the regularization parameter increases, the price of fairness approaches zero, and proportional fairness converges to maximal utility. This bound enables principled selection of the regularization parameter, with theoretical guarantees on the resulting price of fairness. Figure 2 illustrates the bound alongside empirical results on a synthetic distribution.

The proof is deferred to Appendix D. It formulates an optimization problem whose solution is the proportionally fair allocation, derives properties of the optimum via the KKT conditions, and uses them to bound the gap between the optimal and proportionally fair solutions.

5.2 Equal opportunity and max-min fairness

In this section we show that requiring max-min fairness (and equal opportunity for equal base rates) are stricter than demographic parity and in some cases, can cause the utility to approach zero. It suffices to consider two groups, as merging groups only relaxes the constraint.

In Appendix D we state and prove a theorem showing that under tail dominance, max-min fairness is stricter than demographic parity. We show here a gap version of the theorem. First, we define a gap version of tail dominance. Requiring a gap for all $r \geq t_0$ is too restrictive in practice, as high-risk regions may have negligible mass. We therefore impose the gap only on a finite interval.

Definition 5.2. Let $p : \mathcal{X} \to [0,1]$ and S_1, S_2 be disjoint groups. We say that S_1 has (t_0, t_M)-tail dominance over S_2 with gap η if it is t_0 tail dominance, and in addition for all $r \in [t_0, t_M]$,

$$\Pr_{x\sim\mathcal{D}}[p(x) > r | x \in S_1] - \Pr_{x\sim\mathcal{D}}[p(x) > r | x \in S_2] \geq \eta.$$

Theorem 5.3. *Let $\mathcal{D}$ be a $2c$-non-degenerate distribution over $\mathcal{X} \times \{0,1\}$ with groups S_1, S_2. Let $p : \mathcal{X} \to [0,1]$ be α-calibrated, and suppose S_1 has (t_0, t_M)-tail dominance over S_2 with gap η. Assume that for all $r \geq t_0$, $\Pr_{x\sim\mathcal{D}}[p(x) \in [r, r+\beta] \mid x \in S_i] \leq \eta$ and $\Pr_{x\sim\mathcal{D}}[p(x) \geq t_M] \leq \eta$.*

Let $c \in [6\eta, 1]$ satisfy $\Pr[p(x) \geq t_0 \wedge x \in S_i] \geq c$ for $i \in [2]$. Then any randomized threshold τ with capacity c that satisfies max-min fairness on $\mathcal{D}_p$ satisfies

$$\mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(p(x),1) \mid x \in S_1] \;\leq\; \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(p(x),2) \mid x \in S_2] \;-\; \frac{\eta\beta}{t_{\min}}.$$

If τ satisfies max-min fairness up to error δ on $\mathcal{D}$, such that $2\alpha + \delta < \beta\eta$ then

$$\mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(p(x),1) \mid x \in S_1] \;\leq\; \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(p(x),2) \mid x \in S_2] \;-\; \frac{\eta\beta - (2\alpha+\delta)}{t_{\min}}.$$

The proof is deferred to Appendix D. The theorem shows that under a large gap and weak signal in S_2, max-min fairness can allocate a strict majority of resources to S_2, even when its true positive rate is low, leading to a sharp drop in overall performance.

Claim 5.4. Under the conditions of Theorem 5.3, if the groups S_1, S_2 are equal-sized groups, $p(x)$ is continuous and satisfies $\Pr[p(x) \in [r, r+\beta] \mid x \in S_i] \leq \eta/4$. Then the additive price of fairness of requiring max-min fairness with respect to distribution $\mathcal{D}$ with error δ is at least

$$\frac{\eta\beta}{8} + \frac{\beta(\eta\beta - 2\alpha - \delta)}{4t_2} - 4\alpha.$$

We illustrate that the price of fairness can be unbounded under max-min fairness (and equal opportunity). For any $B \in \mathbb{R}$, we construct an instance where the multiplicative price of fairness is at least $B/2$ for sufficiently small capacity.

Example 5.5. Let $B \in \mathbb{R}$ and $0 < c \leq 1/B$. Consider a distribution $\mathcal{D}$ over $\mathcal{X} \times \{0,1\}$ with two groups S_1, S_2, and a perfectly calibrated predictor $p : \mathcal{X} \to [0,1]$ such that for $x \in S_1$, $\Pr[p(x) = 0.1] = c$ and $\Pr[p(x) = 0] = 1 - c$, and for $x \in S_2$, $\Pr[p(x) = 0.1/B] = c$ and $\Pr\left[p(x) = \frac{c(B-1)}{10(1-c)B}\right] = 1 - c$.

The optimal allocation τ' assigns all resources to S_1, yielding $\mathbb{E}[\tau'(p(x), g(x))y] = 0.1c$. Since both groups have identical base rates, max-min fairness and equal opportunity coincide. The fair allocation τ assigns $\frac{c}{B+1}$ to S_1 and the remainder to S_2, yielding $\mathbb{E}[\tau(p(x), g(x))y] = \frac{c}{10(B+1)} + \frac{cB}{10(B+1)B} \leq \frac{2}{10B}\,\mathbb{E}[\tau'(p(x), g(x))y]$.

6 Achievable equal opportunity

We propose a limited-capacity variant of equal opportunity with a bounded price of fairness. Instead of equalizing over all individuals with $y = 1$, we restrict attention to those that can be reached under capacity constraints. This defines the *achievable set*, which depends on both the capacity and the function class.

For $c \in [0,1]$, let $\mathcal{F}_c \subseteq \mathcal{F}$ be functions with capacity c, and define the achievable true positive count

$$\mathcal{A}(\mathcal{F}, c, S_i) := \max_{h \in \mathcal{F}_c} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y \mid x \in S_i].$$

Achievable equal opportunity. A function $f : \mathcal{X} \to [0,1]$ satisfies achievable equal opportunity if for all $i, j \in [m]$,

$$\frac{\mathbb{E}[f(x)y \mid x \in S_i]}{\mathcal{A}(\mathcal{F}, c, S_i)} = \frac{\mathbb{E}[f(x)y \mid x \in S_j]}{\mathcal{A}(\mathcal{F}, c, S_j)}.$$

For $c = 1$, if $\mathcal{F}$ contains the constant function 1, then $\mathcal{A}(\mathcal{F}, c, S_i) = \mathbb{E}[y \mid x \in S_i]$, recovering standard equal opportunity.

Definition 6.1. $\mathcal{F}$ is closed under group-choice post-processing if for any $f_1, \ldots, f_m \in \mathcal{F}$ and any τ, the function $f^*(x) = \tau(f_{g(x)}, g(x))$ belongs to $\mathcal{F}$.

Claim 6.2. Let $\mathcal{F}$ be closed under group-choice post-processing. For any capacity c, let f maximize $\mathbb{E}[f(x)y]$ under achievable equal opportunity, and let h maximize $\mathbb{E}[h(x)y]$ without constraints. Then $\mathbb{E}[f(x)y] \geq \frac{1}{m}\,\mathbb{E}[h(x)y]$.

The proof is deferred to Appendix E, where we also show the bound is tight. We remark that in settings with a large group size imbalance, this definition can be stricter than equal opportunity.

References

[ABFRV22] Georgios Amanatidis, Georgios Birmpas, Aris Filos-Ratsikas, and Alexandros A Voudouris. Fair division of indivisible goods: A survey. *arXiv preprint arXiv:2202.07551*, 2022.

[ANW20] Lovenoor Aulck, Dev Nambi, and Jevin West. Increasing enrollment by optimizing scholarship allocations using machine learning and genetic algorithms. *International Educational Data Mining Society*, 2020.

[BCS25] Kelli A Bird, Benjamin L Castleman, and Yifeng Song. Are algorithms biased in education? exploring racial bias in predicting community college student success. *Journal of Policy Analysis and Management*, 44(2):379–402, 2025.

[BFT11] Dimitris Bertsimas, Vivek F Farias, and Nikolaos Trichakis. The price of fairness. *Operations research*, 59(1):17–31, 2011.

[BGMH24] Michael Barth, Arne Güllich, Brooke N Macnamara, and David Z Hambrick. Quantifying the extent to which junior performance predicts senior performance in olympic sports: a systematic review and meta-analysis. *Sports medicine*, 54(1):95–104, 2024.

[BOT18] Hamoud S Bin-Obaid and Theodore B Trafalis. Fairness in resource allocation: Foundation and applications. In *International Conference on Network Analysis*, pages 3–18. Springer, 2018.

[CB21] Lindsay Cattell and Julie Bruch. Identifying students at risk using prior performance versus a machine learning algorithm. rel 2021-126. *Regional Educational Laboratory Mid-Atlantic*, 2021.

[CDE+26] Arthur Charpentier, Christophe Denis, Romuald Elie, Mohamed Hebiri, and François Hu. Fair regression under localized demographic parity constraints. *arXiv preprint arXiv:2603.25224*, 2026.

[CDPF+17] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pages 797–806, 2017.

[DGK+21] Emily Diana, Wesley Gill, Michael Kearns, Krishnaram Kenthapadi, and Aaron Roth. Minimax group fairness: Algorithms and experiments. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 66–76, 2021.

[DK20] Kate Donahue and Jon Kleinberg. Fairness and utilization in allocating resources with uncertain demand. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 658–668, 2020.

[DOBD+18] Michele Donini, Luca Oneto, Shai Ben-David, John S Shawe-Taylor, and Massimiliano Pontil. Empirical risk minimization under fairness constraints. *Advances in neural information processing systems*, 31, 2018.

[Dud18] Richard M Dudley. *Real analysis and probability*. Chapman and Hall/CRC, 2018.

[EJJ+19] Hadi Elzayn, Shahin Jabbari, Christopher Jung, Michael Kearns, Seth Neel, Aaron Roth, and Zachary Schutzman. Fair algorithms for learning in allocation problems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 170–179, 2019.

[GKR+21] Parikshit Gopalan, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder. Omnipredictors. *arXiv preprint arXiv:2109.05389*, 2021.

[HHYL25] Yutian He, Yankun Huang, Yao Yao, and Qihang Lin. Enforcing fair predicted scores on intervals of percentiles by difference-of-convex constraints. *arXiv preprint arXiv:2505.12530*, 2025.

[HJKRR18] Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In *International Conference on Machine Learning*, pages 1939–1948. PMLR, 2018.

[HNRY23] Lunjia Hu, Inbal Rachel Livni Navon, Omer Reingold, and Chutong Yang. Omnipredictors for constrained optimization. In *International Conference on Machine Learning*, pages 13497–13527. PMLR, 2023.

[HPS16] Moritz Hardt, Eric Price, and Nati Srebro. Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29, 2016.

[LDH26] Naht Sinh Le, Christophe Denis, and Mohamed Hebiri. Demographic parity tails for regression. *arXiv preprint arXiv:2604.02017*, 2026.

[MS94] Michael T Marsh and David A Schilling. Equity measurement in facility location analysis: A review and framework. *European journal of operational research*, 74(1):1–17, 1994.

[MW18] Aditya Krishna Menon and Robert C Williamson. The cost of fairness in binary classification. In *Conference on Fairness, accountability and transparency*, pages 107–118. PMLR, 2018.

[OLP+14] Wlodzimierz Ogryczak, Hanan Luss, Michał Pióro, Dritan Nace, and Artur Tomaszewski. Fair optimization and networks: A survey. *Journal of Applied Mathematics*, 2014(1):612018, 2014.

[Tsi] Boris Tsirelson. Probability theory. Lecture notes, Lecture 4. Tel Aviv University / Hebrew University of Jerusalem.

[WWW+18] Jun-Bo Wang, Junyuan Wang, Yongpeng Wu, Jin-Yuan Wang, Huiling Zhu, Min Lin, and Jiangzhou Wang. A machine learning framework for resource allocation assisted by cloud computing. *IEEE Network*, 32(2):144–151, 2018.

[YNT21] Jiamin Yin, Kee Yuan Ngiam, and Hock Hai Teo. Role of artificial intelligence applications in real-life clinical practice: systematic review. *Journal of medical Internet research*, 23(4):e25759, 2021.

[ZMCY22] Guojun Zhang, Saber Malekmohammadi, Xi Chen, and Yaoliang Yu. Proportional fairness in federated learning. *Transactions on Machine Learning Research*, 2022.

[ZOP24] Anna Zink, Ziad Obermeyer, and Emma Pierson. Race adjustments in clinical algorithms can help correct for racial disparities in data quality. *Proceedings of the National Academy of Sciences*, 121(34):e2402267121, 2024.

A Quantile Functions

Some of the proofs in this paper use quantile functions and their properties. We define the basic definitions and prove some standard properties of the quantile functions below, for completeness.

Definition A.1. Let Q be a distribution over $[0,1]$, the cumulative distribution function (CDF) of Q is the function $F : [0,1] \to [0,1]$ defined by

$$F(t) = \Pr_{u \sim Q}[u \leq t].$$

Definition A.2. Let Q be a distribution over $[0,1]$, the quantile function of Q is the function $q : [0,1] \to [0,1]$ defined by

$$q(\kappa) = \inf \{t \,|F(t) \geq \kappa\}.$$

This is the standard definition of a quantile function, although we remark that there are some cases in the literature that other definitions are sometimes used (where instead of the infimum, the supremum or other expression is used).

We note that for every distribution Q, both the CDF F and the quantile function q are monotonically increasing. In addition, for every $t \in [0,1]$ we have that $q(F(t)) \leq t$ (as we pick q to be the infimum) and $F(q(k)) \geq k$. We further note that for $K > F(t)$ we have that $q(K) > t$, because $q(K) = \inf\{r|F(r) \geq K\}$, since $F(t) < K$, $q(K)$ must be larger than t.

Claim A.3. Let Q be a distribution over $[0,1]$, q be the quantile function of it and F be the commulative distribution function of Q. Then, for every $t < t_M, \in [0,1]$,

$$\mathop{\mathbb{E}}_{u\sim Q}[\mathbf{1}(u > t)u] = \int_{F(t)}^{1} q(r)dr \tag{1}$$

$$\mathop{\mathbb{E}}_{u\sim Q}[\mathbf{1}(t < u \leq t_M)u] = \int_{F(t)}^{F(t_M)} q(r)dr \tag{2}$$

Proof. By definition, the quantile function is monotonically increasing and bounded, and therefore the integrals are well-defined.

We start by proving (1). Let v be the random variable defined by picking $u \sim Q$, then setting $v = 0$ if $u \leq t$ and $v = u$ otherwise. Let Q_v be the distribution of v and q_v be the quantile function of v.

For $r \leq t$, we have that $\Pr_{v\sim Q_v}[v \leq r] = \Pr_{u\sim Q}[u \leq t] = F(t)$. This implies that for all $k < F(t)$ we have that $q_v(k) = \inf\{r \,|\Pr_{v\sim Q_v}[v \leq r] \geq k\} = 0$.

For every $r > t$ we have that $\Pr_{v\sim Q_v}[v \leq r] = \Pr_{u\sim Q}[u \leq r] = F(r)$. Therefore, for $k > F(t)$ we have that $q_v(k) = \inf\{r \,|\Pr_{v\sim Q_v}[v \leq r] \geq k\} = \inf\{r \,|\Pr_{u\sim Q}[u \leq r] \geq k\} = q(k)$.

Combining it together we have:

$$\mathop{\mathbb{E}}_{u\sim Q}[\mathbf{1}(u > t)u] = \mathop{\mathbb{E}}_{v\sim Q_v}[v] = \int_0^1 q_v(r)dr = \int_{F(t)}^1 q(r)dr,$$

where we use the fact that we can express the expected value of any random variable as the integral over its quantiles, see [Dud18, Tsi].

The proof of (2) follows the same approach. We define a new random variable w by sampling $u \sim Q$ and then setting $w = 0$ if $u \leq t$ or $u > t_M$ and $w = u$ otherwise. Let Q_w be the distribution of this new random variable, and q_w be it's quantile function. For $r \leq t$:

$$\Pr_{w\sim Q_w}[w \leq r] = \Pr_{u\sim Q}[u \leq t] + \Pr_{u\sim Q}[u > t_M] = F(t) + (1 - F(t_M)).$$

This implies that for all $k < F(t) + (1 - F(t_M))$ we have $q_w(k) = 0$.

For $t < r \leq t_M$ we have that $\Pr_{w\sim Q_w}[w \leq r] = \Pr_{u\sim Q}[u \leq r] + \Pr_{u\sim Q}[u > t_M] = F(r) + (1 - F(t_M))$, which implies that for $F(t) + (1 - F(t_M)) < k \leq F(t_M) + (1 - F(t_M))$ we have $q_w(k) = q(k - (1 - F(t_M)))$. We note that $F(t_M) + (1 - F(t_M)) = 1$, so we have characterized q_w on its full range.

$$\mathop{\mathbb{E}}_{u\sim Q}[\mathbf{1}(t < u \leq t_M)u] = \mathop{\mathbb{E}}_{w\sim Q_w}[w] = \int_0^1 q_w(k)dk = \int_{F(t)+(1-F(t_M))}^{F(t_M)+(1-F(t_M))} q_w(k)dk = \int_{F(t)}^{F(t_M)} q(r)dr.$$

□

Claim A.4. Let Q be a distribution over $[0,1]$, F its CDF and q its quantile function. Then for every $t \in [0,1], \phi \in [F(t), 1]$,

$$\int_{F(t)}^{\phi} q(r)dr \geq (\phi - F(t))t.$$

Proof. If $\phi = F(t)$ then both sides equal zero and we are done. Therefore, we assume in the rest of the proof that $\phi > F(t)$, and let $\Delta = \phi - F(t) > 0$.

We first show that for all $\epsilon > 0$ we have that $q(F(t) + \epsilon) > t$. This is true since $q(F(t) + \epsilon) = \inf\{r|F(r) \geq F(t) + \epsilon\}$, since $F(t) < F(t) + \epsilon$ and F is monotone, it must be that $q(F(t) + \epsilon) > t$.

Next, we bound the second half of the integral. Denote $\tilde{t} = q(F(t) + 0.5\Delta)$, and from above we know that $\tilde{t} > t$. Then we have:

$$\int_{F(t)+0.5\Delta}^{\phi} q(r)dr \geq 0.5\Delta\tilde{t} > 0.5\Delta t. \tag{3}$$

Assume towards a contradiction that the claim does not hold, that is $\int_{F(t)}^{\phi} q(r)dr < (\phi - F(t))t = \Delta t$. Using (3), this implies that

$$\int_{F(t)}^{F(t)+0.5\Delta} q(r)dr < 0.5\Delta t.$$

Since q is bounded, there must be ϵ such that $q(F(t) + \epsilon) < t$ and we reach a contradiction.

□

B Proof of the preliminaries and definitions section

Proof of Claim 2.4. Fix any $i \in [m]$, and we prove the claim on S_i. Let Q_i be the distribution of $p(x)$ for $x \sim \mathcal{D}$ such that $x \in S_i$. We denote by F_i the CDF of Q_i and by q_i its quantile function (see definitions A.1 and A.2). We define $t_i = q_i(1-c)$ and

$$\gamma_i = \begin{cases} 0 & \Pr_{u\sim Q_i}[u = t_i] = 0, \\ \frac{c-1+F_i(t_i)}{\Pr_{u\sim Q_i}[u=t_i]} & \text{o.w.} \end{cases}$$

We show that $\gamma_i \in [0,1]$: since $t_i = q_i(1-c)$, we have that $F(t_i) \geq 1-c$, implying that $\gamma_i \geq 0$. For the upper bound, $F_i(t_i) = \Pr_{u\sim Q_i}[u \leq t_i] = \Pr_{u\sim Q_i}[u = t_i] + \Pr_{u\sim Q_i}[u < t_i] \leq \Pr_{u\sim Q_i}[u = t_i] + 1 - c$, implying that $\gamma_i \leq 1$. The randomized threshold function τ on S_i is defined by:

$$\tau(x,i) = \begin{cases} 1 & p(x) > t_i, \\ 0 & p(x) < t_i \\ \gamma_i & p(x) = t_i. \end{cases}$$

We show that τ satisfies the capacity requirements:

$$\begin{aligned} \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(p(x),i)|x \in S_i] &= \mathop{\mathbb{E}}_{u\sim Q_i}[\tau(u,i)] \\ &= 1 - F_i(q_i(1-c)) + \Pr_{u\sim Q_i}[u = q_i(1-c)]\gamma_i \\ &= 1 - F_i(q_i(1-c)) + (c - 1 + F_i(q_i(1-c))) = c. \end{aligned}$$

□

Proof of Claim 2.7. Let $\ell : \{0,1\} \times \{0,1\} \to [0,1]$ be any bounded binary loss function with $\ell(0,0) = \ell(1,1) = 0$. Let $h \in \mathcal{F}$ be any function with capacity c. We express the expected loss of h, denoted by $\mathcal{L}_{\mathcal{D}}(h)$:

$$\begin{aligned} \mathcal{L}_{\mathcal{D}}(h) &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)\ell(1,y) + (1-h(x))\ell(0,y)] \\ &= \Pr_{(x,y)\sim\mathcal{D}}[y=1] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(1-h(x))\ell(0,1)|y=1] + \Pr_{(x,y)\sim\mathcal{D}}[y=0] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)\ell(1,0)|y=0] \\ &= \ell(0,1) \Pr_{(x,y)\sim\mathcal{D}}[y=1] \left(1 - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=1]\right) + \ell(1,0) \Pr_{(x,y)\sim\mathcal{D}}[y=0] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=0] \end{aligned}$$

The function h has capacity exactly c:

$$\Pr_{(x,y)\sim\mathcal{D}}[y=0] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=0] = c - \Pr_{(x,y)\sim\mathcal{D}}[y=1] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=1].$$

We use this expression in the expected loss:

$$\begin{aligned} \mathcal{L}_{\mathcal{D}}(h) &= \ell(0,1) \Pr_{(x,y)\sim\mathcal{D}}[y=1] \left(1 - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=1]\right) + \ell(1,0) \left(c - \Pr_{(x,y)\sim\mathcal{D}}[y=1] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|y=1]\right) \\ &= \ell(0,1) \Pr_{(x,y)\sim\mathcal{D}}[y=1] + \ell(1,0)c - (\ell(1,0) + \ell(0,1)) \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y]. \end{aligned}$$

The only term in $\mathcal{L}_{\mathcal{D}}(h)$ that depends on h is $-(\ell(1,0) + \ell(0,1)) \mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y]$. Therefore, any function $f \in \mathcal{F}$ that maximizes $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y]$ is also the function that minimizes $\mathcal{L}_{\mathcal{D}}(h)$. □

In the next claim we formalize the normalization requirement from Definition 3.2, that a uniform splitting of a group would not change the fairness requirement. This claim proves that the normalization we chose is the appropriate one for the fairness definition.

Claim B.1. Let $S_1, \ldots, S_m$ be a partition of $\mathcal{X}$. For any $\eta \in (0,1)$, let $S_1^{(1)}, S_1^{(2)}$ be the groups created by splitting S_1 into two random subgroups of sizes $\eta, (1-\eta)$ respectively. Then for any function $f : \mathcal{X} \to [0,1]$,

$$\mathbf{Prop}_{\mathcal{D},S_1,S_2,\ldots,S_m}(f) = \mathbf{Prop}_{\mathcal{D},S_1^{(1)},S_1^{(2)},S_2,\ldots,S_m}(f)$$

Proof. The groups $S_1^{(1)}, S_1^{(2)}$ are random subsets of S_1, so for every function $f : \mathcal{X} \to [0,1]$

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D},S_1^{(1)},S_1^{(2)}}\left[f(x)y|x \in S_1^{(1)}\right] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D},S_1^{(1)},S_1^{(2)}}\left[f(x)y|x \in S_1^{(2)}\right] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}\left[f(x)y|x \in S_1\right],$$

where the randomness is also over the choice of $S_1^{(1)}, S_1^{(2)}$. By definition,

$$\begin{aligned}
\mathbf{Prop}_{\mathcal{D},S_1^{(1)},S_1^{(2)},S_2,\ldots,S_m}(f) = &\Pr_{(x,y)\sim\mathcal{D}}\left[x \in S_1^{(1)}\right]\log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}\left[f(x)y|x \in S_1^{(1)}\right] \\
&+ \Pr_{(x,y)\sim\mathcal{D}}\left[x \in S_1^{(2)}\right]\log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}\left[f(x)y|x \in S_1^{(2)}\right] \\
&+ \sum_{i=2}^{m} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]\log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] \\
= &\left(\Pr_{(x,y)\sim\mathcal{D}}\left[x \in S_1^{(1)}\right] + \Pr_{(x,y)\sim\mathcal{D}}\left[x \in S_1^{(2)}\right]\right)\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_1] \\
&+ \sum_{i=2}^{m} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]\log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] \\
= &\mathbf{Prop}_{\mathcal{D},S_1,S_2,\ldots,S_m}(f)
\end{aligned}$$

□

C Proof of Section 4

Proof of Claim 4.3. Assume towards a contradiction that we are in a c non-degenerated distribution $\mathcal{D}$, and that the function $f \in \mathcal{F}$ satisfying max-min fairness does not have equal true positive count between groups. We find a new function $h \in \mathcal{F}$ that is a post-processing of f and has a higher max-min value, contradicting the optimality of f.

Let $i_1, \ldots, i_k \in [m]$ be all of the group in $[m]$ with a minimal true positive count, i.e. minimal $\mu_{\min} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_j}]$ (it is possible that $k = 1$ and there is a single group with a minimal value). According to our assumption of unequal true positive count between groups, there exists at least one $i_{\max} \in [m]$ with $\mu_{\max} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_{\max}}] > \mu_{\min}$. Set

$$\begin{aligned}
\delta &= \min\left\{\frac{1}{2}\left(\mu_{\max} - \mu_{\min}\right), \Pr_{(x,y)\sim\mathcal{D}}\left[x \in \cup_{j\in[k]}S_{i_j}, f(x) \le 1-\mu_{\min}\right]\mu_{\min}\right\} \\
\epsilon &= \delta\frac{\Pr_{(x,y)\sim\mathcal{D}}[x \in S_{i_{\max}}, f(x) \ge \delta]}{\Pr_{(x,y)\sim\mathcal{D}}\left[x \in \cup_{j\in[k]}S_{i_j}, f(x) \le 1-\mu_{\min}\right]}
\end{aligned}$$

We note that $\mu_{\min} < c < 0.5$ and so $1 - \mu_{\min} > \mu_{\min}$. Combining with the fact that $f(x) \ge 0$ we get that $\Pr_{(x,y)\sim\mathcal{D}}\left[x \in \cup_{j\in[k]}S_{i_j}, f(x) < 1-\mu_{min}\right] > 0$ and ϵ is well-defined.

We define a new function h that is a post-processing of f by $h(x) = \tau(f(x), g(x))$ for:

$$\tau(r,i) = \begin{cases} r-\delta & \text{if } i = i_{\max} \text{ and } r \ge \delta \\ r+\epsilon & \text{if } i \in \{i_1,\ldots,i_k\} \text{ and } r \le 1-\mu_{\min} \\ r & \text{o.w..} \end{cases}$$

That is, we increase f on $S_{i_j}, \ldots, S_{i_j}$ and decrease its values on $S_{i_{\max}}$. The condition of $r \geq \delta$ is to ensure that the output of h is in $[0,1]$ (as $\delta \in (0,1]$). For ϵ, we need to show that $\epsilon \in [0, \mu_{\min}]$:

$$\epsilon \leq \frac{\delta}{\Pr_{(x,y)\sim\mathcal{D}}\left[x \in \cup_{j\in[k]} S_{i_j}, f(x) < \mu_{max}\right]} \leq \mu_{\min}$$

by the definition of δ, and therefore the output of h is in $[0,1]$.

We show next that h has capacity exactly c:

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)] &= \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|x \in S_i] \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)] - \delta \Pr_{(x,y)\sim\mathcal{D}}[x \in S_{i_{\max}}, f(x) \geq \delta] + \epsilon \Pr_{(x,y)\sim\mathcal{D}}\left[x \in \cup_{j\in[k]} S_{i_j}, f(x) \leq 1 - \mu_{\min}\right] \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)].
\end{aligned}$$

To finish the proof, we show that h has higher max-min value than the original f, contradicting it's optimality. We show it by going over all groups $i \in [m]$:

- For $i_{\max}$: $\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_{i_{\max}}] \geq \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_{\max}}] - \delta > \mu_{\min}$, as $\delta < \mu_{\max} - \mu_{\min}$.
- For $i \notin \{i_1, \ldots, i_k\} \cup \{i_{\min}\}$: by definition $\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i] = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] > \mu_{\min}$.
- For $i \in \{i_1, \ldots, i_k\}$: Let f_{in} be defined by $f_{in}(x) = \mathbf{1}(f(x) > 1 - \mu_{\min})$, where $\mathbf{1}$ is the indicator function. Then f_{in} is a post-processing of f and therefore in $\mathcal{F}$. In addition,$c = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)] \geq (1 - \Pr_{(x,y)\sim\mathcal{D}}[f(x) \leq 1 - \mu_{\min}])(1 - \mu_{\min})$ implying that $\Pr_{(x,y)\sim\mathcal{D}}[f_{in}(x) = 1] = \Pr_{(x,y)\sim\mathcal{D}}[f(x) > 1 - \mu_{\min}] \leq 1 - \frac{c}{1-\mu_{\min}} \leq 2c$. From the non-degeneracy condition, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[y|x \in S_{i_j}, f_{in}(x) = 0] > 0$. From this we can bound:

  $$\begin{aligned}
  \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_{i_j}] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_j}] + \epsilon \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[y\mathbf{1}(f_{in}(x) = 1)|x \in S_{i_j}] \\
  &> \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_j}]
  \end{aligned}$$

  where the inequality is because of the non degenerate condition, and the probability of $f(x) \leq 1 - \mu_{\min}$ for $x \in S_{i_j}$ is strictly positive.

□

We show next that the set of all threshold functions of a predictor equalizes the group utilities.

Claim C.1. For $c < 0.5$, let $\mathcal{D}$ be $2c$-non-degenerate with respect to groups $S_1, \ldots, S_m$, and let $\mathcal{F}$ be the set of all randomizes group threshold functions. Then any $f \in \mathcal{F}$ that satisfies max-min fairness with capacity c over a calibrated predictor y equalizes group utilities, i.e., for all $i, j \in [m]$,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y \mid x \in S_i] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y \mid x \in S_j].$$

Proof. Assume towards a contradiction that there exists a randomized group threshold function $f \in \mathcal{F}$ satisfying max-min fairness does not equalizes the true positive count. Let $i_1, \ldots, i_m$ be the indices of all groups with minimal true positive count, and define $\mu_{\min} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_j}]$. Let $c_1, \ldots, c_k$ be the capacity of f restricted to groups $i_1, \ldots, i_k$. That is, $c_j = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_{i_j}]$. From our non-degeneracy assumption, $c_j < 1$ for all j.

According to our assumption of unequal true positive count between groups, there exists at least one $i_{\max} \in [m]$ with $\mu_{\max} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_{\max}}] > \mu_{\min}$. Let $c_{\max} = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_{i_{\max}}]$.

Let $\delta = (\mu_{\max} - \mu_{\min})/2$. If there exists j such that

$$c_j + \frac{\delta \Pr[x \in S_{i_{\max}}]}{k \Pr[x \in S_{i_j}]} > 1,$$

then we reduce δ to $(1 - c_j)k \Pr[x \in S_{i_j}] / \Pr[x \in S_{i_{\max}}]$, which is a strictly positive value, since from our non-dejeneracy assumption $c_j < 1$.

Then, we define a new randomized threshold function f' as the function satisfying:

- For $x \notin S_{i_1} \cup \cdots \cup S_{i_k} \cup S_{i_{\max}}$, $f'(x) = f(x)$.
- For $x \in S_{i_{\max}}$, f' is the randomized threshold function with capacity $c_{\max} - \delta$ on $S_{i_{\max}}$, promised from Claim 2.4. We note that $c_{\max} - \delta > 0$ since $c_{\max} > \mu_{\max}$.
- For $x \in S_{i_j}, j \in [k]$, f' is the randomizes threshold function with capacity $c_j + \frac{\delta \Pr[x \in S_{i_{\max}}]}{k \Pr[x \in S_{i_j}]}$, promised from Claim 2.4. We note that from our choice of δ, the capacity is always at most 1.

By definition, if f is a randomized threshold function, then so is f'. We show that they also have the same capacity:

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f'(x)] &= \sum_{i\in[m]} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f'(x)|x \in S_i] \\
&= \sum_{i\notin\{i_1,\ldots,i_k\}\cup\{i_{\max}\}} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_i] + \Pr[x \in S_{i_{\max}}](c_{\max} - \delta) \\
&\quad + \sum_{j\in[k]} \Pr[x \in S_{i_j}] \left(c_j - \frac{\delta \Pr[x \in S_{i_{\max}}]}{k \Pr[x \in S_{i_j}]} \right) \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)] - \delta \Pr[x \in S_{i_{\max}}] + k \frac{\delta \Pr[x \in S_{i_{\max}}]}{k} \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)].
\end{aligned}$$

We show next that f' has strictly larger max-min value, contradicting our assumption.

- For $i_{\max}$: $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f'(x)y|x \in S_{i_{\max}}] \geq \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i_{\max}}] - \delta > \mu_{\min}$, as $\delta < \mu_{\max} - \mu_{\min}$.
- For $i \notin \{i_1, \ldots, i_k\} \cup \{i_{\min}\}$: by definition $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f'(x)y|x \in S_i] = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] > \mu_{\min}$.
- For $i \in \{i_1, \ldots, i_k\}$: for every i_j let $f_{in}(x) = \mathbf{1}(f(x) = 1)$, then it's capacity is at most c, and by the non-degeneracy condition and the fact that p is calibrated, there is a non-zero probability for $y = 1$ on those x in which we increase f.

□

Proof of Lemma 4.4. Let $f \in \mathcal{F}$ be the function that satisfies max-min fairness with respect to $S_1, \ldots, S_m$, and let $h \in \mathcal{F}$ be the function satisfying equal opportunity with respect to these groups and maximizing $\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y]$, as in the Lemma statement. We prove the claim by showing the both f, h satisfy both conditions, i.e. that f satisfies equal opportunity and maximizes $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y]$, and that h satisfies max-min fairness.

First, we show that f also satisfies the equal opportunity constraints. From the conditions of the claim, we have f equalizes the group utility, i.e. $i, j \in [m]$: $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] = \mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_j]$, and we denote this value by μ_f. Since we have equal base rates, for all $i, j \in [m]$

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)|y = 1, x \in S_i] = \frac{\mu_f}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[y = 1|x \in S_i]} = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)|y = 1, x \in S_j].$$

Which means that f satisfies equal opportunity.

For h, the equal opportunity constraints together with the equal base rates implies that for all $i, j \in [m]$:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[y|x \in S_i] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)|, y = 1x \in S_i] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_j].$$

We denote this value by $\mu_h = \mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i]$. We note that by the max-min definition, we have that $\min_{i\in[m]} \mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i] = \mu_h$.

Using this notation

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y] = \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i] = \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]\mu_h = \mu_h. \tag{4}$$

Using an identical argument on f we have that, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y] = \mu_f$.

Since $h \in \mathcal{F}$ is the function in $\mathcal{F}$ that satisfies equal opportunity and maximizes $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y]$ and f satisfies equal opportunity, we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y] = \mu_h \geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y] = \mu_f.$$

The function f is satisfies max-min fairness, therefore

$$\min_{i\in[m]} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_i] = \mu_h \leq \min_{i\in[m]} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] = \mu_f.$$

We conclude that $\mu_f = \mu_h$, and that both functions satisfy both requirements. □

Proof of Lemma 4.6. We write the proportional fairness optimization expression explicitly:

$$\begin{aligned}
\mathbf{Prop}_{\mathcal{D},\mathcal{S}}(f) &= \sum_{i\in[m]} \Pr_{x\sim\mathcal{D}}[x \in S_i] \log u_i(f) \\
&= \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \log \frac{\Pr_{x\sim\mathcal{D}}[x \in S_i]u_i(f)}{\Pr_{x\sim\mathcal{D}}[x \in S_i]} \\
&= \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \log \frac{\Pr_{x\sim\mathcal{D}}[x \in S_i]u_i(f)U(f)}{\Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]U(f)} \\
&= \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \left(\log \frac{\Pr_{x\sim\mathcal{D}}[x \in S_i]u_i(f)}{\Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]U(f)} + \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y]\right) \\
&= - D_{KL}(Q_{\mathcal{S}}, Q_f) + \log U(f)
\end{aligned}$$

Where the minus sign comes from flipping the denominator and numerator to get the divergence. □

Proof of Claim 4.7. From Claim 4.3, a function f satisfying max-min fairness requirement for a non-degenerate distribution satisfies that for all $i \neq j$:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_j] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y].$$

Such f satisfies for every $i \in [m]$:

$$\Pr_{j\sim Q_f}[j = i] = \frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y\mathbf{1}(x \in S_i)]}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y]} = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(x \in S_i)] = \Pr_{j\sim Q_{\mathcal{S}}}[j = i].$$

Which implies that $D_{KL}(Q_{\mathcal{S}}, Q_f) = 0$. □

D Proof of Section 5

D.1 Proportional fairness

We start by bounding the difference between the true positive rate of the simulated distribution $\mathcal{D}_p$ and the distribution $\mathcal{D}$.

Claim D.1. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$ where $\mathcal{X}$ is divided into groups $S_1, \ldots, S_m$. Let $p : \mathcal{X} \to [0,1]$ be a predictor that is α-calibrated on S_1, S_2, satisfying for all $i \in [m]$

$$\mathop{\mathbb{E}}_{x' \sim \mathcal{D}, x' \in S_i} \left[\left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[y | x \in S_i, p(x) = p(x')] - p(x') \right| \right] \leq \alpha. \tag{5}$$

Then, for every randomized group threshold function $\tau : [0,1] \times [m] \to [0,1]$ and every group i, if $f : \mathcal{X} \to [0,1]$ is defined by $f(x) = \tau(p(x), g(x))$ for all x, then

$$\left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)y | x \in S_i] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)p(x) | x \in S_i] \right| \leq \alpha. \tag{6}$$

Proof. Fix some $i \in [m]$ and let t_i, γ_i be the threshold and value at the threshold of τ. That is, $\tau(t_i, i) = \gamma_i$ and for all $r > t_i$, $\tau(r, i) = 1$. By definition,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)y | x \in S_i] = \gamma_i \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_i)y | x \in S_i] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t_i)y | x \in S_i].$$

Then we can bound the error in approximation, using the triangle inequality

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)y | x \in S_i] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)p(x) | x \in S_i] \right| \\
\leq &\gamma_i \left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_i)(y - p(x)) | x \in S_i] \right| + \left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t_i)(y - p(x)) | x \in S_i] \right| \\
\leq &\mathop{\mathbb{E}}_{x' \sim \mathcal{D}, x' \in S_i} \left[\left| \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[y | x \in S_i, p(x) = p(x')] - p(x') \right| | p(x) \geq t_i \right] \leq \alpha.
\end{aligned}$$

□

Proof of Theorem 5.1. Let $\mathcal{D}, p$ as in the theorem statement. Let $\tau : [0,1] \times [m] \to [0,1]$ be the post-processing of p that satisfies proportional fairness. Denote by $t_i, i \in [m]$ its thresholds.

By definition, the post-processing of p that satisfies propositional fairness and has capacity c is give by $\arg\max_{\tau \text{ capacity } c} \mathbf{Prop}_{\mathcal{D}_p,S}(\tau)$. where the proportional fairness expression is

$$\begin{aligned}
\mathbf{Prop}_{\mathcal{D},S}(\tau) &= -D_{KL}(Q_{\mathcal{S}}, Q_f) + \lambda \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(x)y] \\
&= \sum_{i \in [m]} \Pr_{(x,y)\sim\mathcal{D}_p}[x \in S_i] \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(x)y | x \in S_i] + (\lambda - 1) \log \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(x)y].
\end{aligned}$$

By the definition of the simulated distribution, this is equivalent to

$$\mathbf{Prop}_{\mathcal{D}_p,S}(\tau) = \sum_{i \in [m]} \Pr_{x\sim\mathcal{D}}[x \in S_i] \log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(x)p(x) | x \in S_i] + (\lambda - 1) \log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau(x)p(x)].$$

For each i, denote $c_i = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)\mathbf{1}(x \in S_i)]$, this is the capacity group i receives from c. We note that c_i is not the conditional capacity, but rather the part of c allocated to group i taking into account its size. Denote by C the ordered set of the capacities: $C = (c_1, \ldots, c_m)$. For every $C = (c_1, \ldots, c_m)$ such that $c_i \leq \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]$, let τ_C be the randomized threshold function such that for all i, $c_i = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(p(x), i)\mathbf{1}(x \in S_i)]$. By Claim 2.4, we can always find such τ_C.

Using this notation, we can express the post processing τ satisfying proportional fairness as $\tau = \tau_C$ for C that is the solution to the following optimization problem:

$$\max_{c_1,\ldots,c_m} \mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) \tag{7}$$

$$\forall i, \quad 0 \leq c_i \leq \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i] \tag{8}$$

$$\sum_{i \in [m]} c_i \leq c. \tag{9}$$

By Claim D.2, this maximization problem has concave optimization goal. We can also see that all of the constraints are linear. Therefore, in order to understand the properties of the solution, we analyze the KKT conditions for this optimization problem. We note that there is a strong duality, since the problem is a concave maximization problem and there is a point strictly inside the feasible set, as we can simply choose to $c_i = \min\{\Pr[x \in S_i]/2, c/(2m)\}$. Therefore, from Slater's condition, we have strong duality.

Let $\eta, i \in [m]$ be the variables of the constraint $c_i \leq \Pr[x \in S_i]$ and β_i the dual variables of $c_i \geq 0$ and Γ be the dual variable of the capacity constraint. The KKT condition is:

$$\nabla_C \mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) + \sum_{i\in[m]} \eta_i \nabla_C c_i + \beta_i \nabla_C(-c_i) + \Gamma \nabla_C \left(\sum_i c_i - c\right) = 0$$

This means that for all $i \in [m]$,

$$\frac{\partial}{\partial c_i}\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) + \eta_i - \beta_i + \Gamma = 0. \tag{10}$$

We show next that the capacity constraint is tight at the maximum point, implying that $\Gamma \neq 0$. Assume towards contradiction that the capacity constraint is not tight, i.e. that $\sum_i c_i < c$. Then let $\tilde{C} = (c_1 + \epsilon, c_2, \ldots, c_m)$ for some $0 \leq \epsilon \leq c - \sum_i c_i$. Since we are in a non-degenerate setting, we have that $c_1 + \epsilon < \Pr[x \in S_1]$. Therefore, $\tilde{C}$ satisfies the constraints. Since we are in a non-degenerate setting, increasing c_1 increases $\mathbf{Prop}_{\mathcal{D}_p,S}$, as we increase $\mathbb{E}[\tau(p(x),1)|x \in S_1]$ and do not change any of the other expressions. Therefore, we conclude that the capacity constraint is tight at the maximum point.

Using similar arguments, we show that none of the other constraints are tight at the maximum. The constraints $c_i < \Pr[x \in S_i]$ cannot be tight in a non-degenerate setting (as c is not large enough) and can in fact be discarded. The constraint $c_i \geq 0$ cannot be tight at the maximum, because in this case the optimization goal contains a logarithm of 0, which cannot be the maximum point. Therefore, we conclude that for all i, $\eta_i, \beta_i = 0$.

That is, for all i, j we have that

$$\frac{\partial}{\partial c_i}\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) = \frac{\partial}{\partial c_j}\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) = -\Gamma \tag{11}$$

Let $C = (c_1, \ldots, c_m)$ be the solution to the maximization problem and let $\tau = \tau_C$ be the post-processing using capacities C. We use (11) to characterize the post-processing τ achieving proportional fairness, and use it to bound the price of fairness.

Let τ' be the post-processing maximizing the utility with capacity c. From Claim D.7, we can assume that τ' has a single threshold t' to all groups $S_1, \ldots, S_m$. Let $c'_1, \ldots, c'_m$ be the capacities that τ' allocates to each group.

We reorder the groups such that for all $i \leq k$ we have that $c'_i > c_i$, and for all $j > k$ we have $c'_j \leq c_j$. Since τ' uses a single threshold t', we have that for all $i \leq k$, $t_i(c_i) \geq t'$. For each $x \in S_i$ such that $\tau(p(x), i) < \tau'(p(x), i)$ we also have that $p(x) \leq t_i(c_i)$. Similarly, for all $j > k$, $t_j(c_j) \leq t'$ and for all $x \in S_j$ such that $\tau(p(x), i) > \tau'(p(x), i)$ we have that $p(x) \geq t_j(c_j)$.

Using the above inequalities, we can bound the gap in the true positive count of τ, τ' as a function of the thresholds.

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(\tau'(p(x), g(x)) - \tau(p(x), g(x)))p(x)] &= \sum_{i\in[m]} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(x \in S_i)(\tau'(p(x), i) - \tau(p(x), i))p(x)] \\ &\leq \sum_{i\in[k]} (c'_i - c_i)t_i(c_i) - \sum_{k<j\leq m} (c_j - c'_j)t_j(c_j)\end{aligned} \tag{12}$$

We use (11) to lower bound $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)]$ in terms of the differences $t_i(c_i) - t_j(c_j)$. Let $i \le k$ and $j > k$ be two groups, then from (11) and Claim D.3,

$$\frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i]} + \frac{(\lambda-1)t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)]} = \frac{t_j(c_j)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j]} + \frac{(\lambda-1)t_j(c_j)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),p(x))p(x)]}. \tag{13}$$

We note that since $t_i(c_i) \ge t' \ge t_j(c_j)$, it must be that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)|x \in S_i] \ge \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)|x \in S_j]$. We reorder the groups such that for all $i \le k'$ we have that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)|x \in S_i] \ge \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)]$, and from above, we know that all of these i we have $c'_i \ge c_i$. Let $i \le k'$. For such i, we replace $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)]$ by $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)|x \in S_i]$, which is equivalent to subtracting $(\lambda-1)(1/\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)] - 1/\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(x)p(x)|x \in S_i])$ times $t_i(c_i)$ from the left hand side and times $t_j(c_j)$ from the right hand side. We get the inequality:

$$\frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i]} + \frac{(\lambda-1)t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i]} \le \frac{t_j(c_j)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j]} + \frac{(\lambda-1)t_j(c_j)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i]}.$$

By some algebra we have that

$$\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i] \ge (\lambda t_i(c_i) - (\lambda-1)t_j(c_j))\frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j]}{t_j(c_j)}$$

For $k' \le i \le k$ we have a weaker inequality:

$$\frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i]} \le \frac{t_j(c_j)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j]}.$$

Note that all inequalities holds for all $j \ge k$, and therefore also over a distribution over them. We also have that for every j, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j] \ge \frac{t_j(c_j)c_j}{\Pr_{(x,y)\sim\mathcal{D}}[x\in S_j]}$, as the probability of $x \in S_j$ to have a positive τ is $c_j/\Pr_{(x,y)\sim\mathcal{D}}[x \in S_j]$, and each such x has $p(x) \ge t_j(c_j)$. That is, for every $i \le k'$ we have

$$\begin{aligned}
&\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i] \\
&\quad\ge \sum_{j\ge k} \frac{\Pr_{(x,j)}[x \in S_j]}{\Pr_{(x,j)}[x \in \cup_{j'\ge k} S_{j'}]}(\lambda t_i(c_i) - (\lambda-1)t_j(c_j))\frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x \in S_j]}{t_j(c_j)} \\
&\quad\ge \sum_{j\ge k} \frac{c_j}{\Pr_{(x,j)}[x \in \cup_{j'\ge k} S_{j'}]}(\lambda t_i(c_i) - (\lambda-1)t_j(c_j))
\end{aligned}$$

Similarly, for each $k' \le i \le k$ we have that

$$\begin{aligned}
\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i] &\ge \sum_{j\ge k} \frac{\Pr_{(x,j)}[x \in S_j]}{\Pr_{(x,j)}[x \in \cup_{j'\ge k} S_{j'}]} \frac{t_i(c_i)\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_j]}{t_j(c_j)} \\
&\ge \sum_{j\ge k} \frac{c_j}{\Pr_{(x,j)}[x \in \cup_{j'\ge k} S_{j'}]} t_i(c_i)
\end{aligned}$$

Combining everything together, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] &= \sum_{i\le k'} \Pr_{(x,y)\sim\mathcal{D}}[x\in S_i] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x\in S_i] \\
&\quad + \sum_{k'\le i\le k} \Pr_{(x,y)\sim\mathcal{D}}[x\in S_i] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x\in S_i] \\
&\quad + \sum_{j>k} \Pr_{(x,y)\sim\mathcal{D}}[x\in S_j] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x\in S_j] \\
&\ge \sum_{i\le k', j>k} \frac{\Pr_{(x,y)\sim\mathcal{D}}[x\in S_i]}{\Pr_{(x,j)}[x\in \cup_{j'\ge k} S_{j'}]} c_j(\lambda t_i(c_i) - (\lambda-1)t_j(c_j)) \\
&\quad + \sum_{k'\le i\le k, j>k} \frac{\Pr_{(x,y)\sim\mathcal{D}}[x\in S_i]}{\Pr_{(x,j)}[x\in \cup_{j'\ge k} S_{j'}]} c_j t_i(c_i)
\end{aligned}$$

Let S_M be the dominant group promised from the theorem, after reordering. We have that $\Pr[x\in S_M]=\mu$. We start by showing that $t_M(c_M)\ge t_j(c_j)$ for every $j\in[m]$. From (13):

$$\begin{aligned}
t_M(c_M)\left(\frac{1}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),M)p(x)|x\in S_M]} + \frac{(\lambda-1)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)]}\right) = \\
t_j(c_j)\left(\frac{1}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x\in S_j]} + \frac{(\lambda-1)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)]}\right). \qquad (14)
\end{aligned}$$

Assume towards a contradiction that $t_M(c_M) < t_i(c_i)$, then from the dominance of S_M and since we increase $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),M)p(x)|x\in S_M]$ by reducing the threshold, we have that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),M)p(x)|x\in S_M]\ge \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),j)p(x)|x\in S_j]$ for all $j\in[m]$. This contradicts (14), as both terms we multiply are smaller in M than in j. Therefore, we conclude that $t_M(c_M)\ge t_j(c_j)$.

Since $t_M(c_M)$ is maximal, in order to satisfy (14) it must be that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x),M)p(x)|x\in S_M]$ is maximal as well, and in particular

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),M)p(x)|x\in S_M] \ge \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)].$$

Therefore, after our reordering we have that $M\le k'$.

By using S_M we can lower bound the true positive count of τ. We have that $\Pr[x\in S_M]=\mu$.

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] \ge \frac{\mu}{1-\mu} c_{gap} t_M(c_M) + \frac{\mu(\lambda-1)}{1-\mu}\sum_j c_j(t_M(c_M)-t_j(c_j))$$

We lower bound the gap, from (12) we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(\tau'(p(x),g(x)) - \tau(p(x),g(x)))p(x)] &\le \sum_{i\in[k]}(c_i'-c_i)t_i(c_i) - \sum_{k<j\le m}(c_j-c_j')t_j(c_j) \\
&\le \sum_{i\in[k]}(c_i'-c_i)t_M(c_M) - \sum_{k<j\le m}(c_j-c_j')t_j(c_j) \\
&\le c_{gap}t_M(c_M) - \sum_{k<j\le m}(c_j-c_j')t_j(c_j) \\
&\le \sum_{k<j\le m}(c_j-c_j')(t_M(c_M)-t_j(c_j)) \\
&\le \sum_{k<j\le m} c_j(t_M(c_M)-t_j(c_j)).
\end{aligned}$$

Combining the two inequalities, we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(\tau'(p(x),g(x)) - \tau(p(x),g(x)))p(x)] \le \frac{1-\mu}{\mu\lambda} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)]$$

That is,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x),g(x))p(x)] \leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)]\left(1+\frac{1-\mu}{\mu\lambda}\right).$$

We can see that as the group S_m is larger, and as the regularization factor is larger the price of fairness goes to zero. We note that we only take into account regularization factors $\lambda > 1$, as our goal is to weaken the proportional fairness definition, that has $\lambda = 1$, and not to strengthen it. We can also see that when $\lambda = 0$, which corresponds to max-min fairness, we have no guarantee.

We are left with handling the calibration error, as in the proof so far we assumed the true positive count is $\tau'(p(x), g(x))p(x)$, i.e that y is distributed according to our predictor. The predictor p a calibration error of at most α, so from Claim D.1, the post-processing that maximizes the utility has different true positive count from τ' by at most 2α. If we enforce proportional fairness on the simulated distribution, then the gap between the simulated distribution and the true positive count of the proportional fairness is also at most 2α, from Claim D.1.

If we have a post-processing τ that satisfies proportional fairness on distribution $\mathcal{D}$, then □

Claim D.2. In the setting of Theorem 5.1, the function $\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C)$ is a concave function in $c_1,\ldots,c_m$.

Proof. We show that the function is concave by showing that for every $C = (c_1,\ldots,c_m)$ and $\tilde{C} = (\tilde{c}_1,\ldots,\tilde{c}_m)$,

$$\mathbf{Prop}(\tau_C) - \mathbf{Prop}(\tau_{\tilde{C}}) \leq \sum_{i\in[m]} \frac{\partial}{\partial c_i}\mathbf{Prop}(\tau_{\tilde{C}})\cdot(c_i-\tilde{c}_i),$$

We start by writing the difference explicitly:

$$\begin{aligned}
&\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) - \mathbf{Prop}_{\mathcal{D}_p,S}(\tau_{\tilde{C}}) \\
&= \sum_{i\in[m]} \Pr_{x\sim\mathcal{D}}[x\in S_i]\left(\log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau_C(p(x,i))p(x)|x\in S_i] - \log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x,i))p(x)|x\in S_i]\right) \\
&+ (\lambda-1)\left(\log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau_C(p(x),g(x))p(x)] - \log \mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),g(x))p(x)]\right) \\
&= \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x\in S_i]\log\left(\frac{\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(p(x),i)p(x)|x\in S_i]}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) \\
&+ (\lambda-1)\log\left(\frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)y]}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_{\tilde{C}}(x)y]}\right)
\end{aligned}$$

We write the partial derivatives explicitly, using Claim D.3:

$$\sum_{i\in[m]} \frac{\partial}{\partial c_i}\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_{\tilde{C}})\cdot(c_i-\tilde{c}_i) = \sum_{i\in[m]} \frac{t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]} + \frac{(\lambda-1)t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),g(x))p(x)]}$$

For all $c_i, \tilde{c}_i$ we have that:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau_C(p(x),i)p(x)|x\in S_i] \leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i] + t_i(\tilde{c}_i)(c_i-\tilde{c}_i)$$

This is because if $c_i > \tilde{c}_i$, then we have that $t_i(c_i) \leq t_i(\tilde{c}_i)$ and the extra capacity $c_i - \tilde{c}_i$ has $p(x) \leq t_i(\tilde{c}_i)$ for all of the x's. If $c_i < \tilde{c}_i$, then the expression $t_i(\tilde{c}_i)(c_i - \tilde{c}_i)$ is negative. In this case $t_i(c_i) \geq t_i(\tilde{c}_i)$, and the extra capacity $\tilde{c}_i - c_i$ has $p(x) \geq t_i(\tilde{c}_i)$, and the same bound holds.

Using this inequality, we have that for all i,

$$\begin{aligned}
\log\left(\frac{\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(p(x),i)p(x)|x\in S_i]}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) &\leq \log\left(1+\frac{t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) \\
&\leq \frac{t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}. \qquad (15)
\end{aligned}$$

In the last inequality we use the fact that $\log(1+z) \leq z$ for all $z > -1$.

Equation (15) is enough to show that if $\lambda = 1$, then $\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C)$ is concave in C. To finish the proof for $\lambda > 1$, we use the same argument.

$$\begin{aligned}
&\log\left(\frac{\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(p(x),g(x))p(x)]}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),g(x))p(x)]}\right) \\
&= \log\left(\frac{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\,\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(p(x),i)p(x)|x\in S_i]}{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\,\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) \\
&\leq \log\left(\frac{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\,(\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]+t_i(\tilde{c}_i)(c_i-\tilde{c}_i))}{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\,\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) \\
&\leq \log\left(1+\frac{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\,\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),i)p(x)|x\in S_i]}\right) \\
&\leq \frac{\sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]t_i(\tilde{c}_i)(c_i-\tilde{c}_i)}{\mathbb{E}_{x\sim\mathcal{D}}[\tau_{\tilde{C}}(p(x),g(x))p(x)]}
\end{aligned}$$

Combining both inequalities we get that

$$\mathbf{Prop}(\tau_C) - \mathbf{Prop}(\tau_{\tilde{C}}) \leq \sum_{i\in[m]}\frac{\partial}{\partial c_i}\mathbf{Prop}(\tau_{\tilde{C}})\cdot(c_i-\tilde{c}_i),$$

and our optimization goal is concave. □

Claim D.3. Let $\mathcal{D}$ be a distribution over $\mathcal{X}\times\{0,1\}$ and $p:\mathcal{X}\to[0,1]$ be a predictor. Let $C=(c_1,\ldots,c_m)$ be a vector such that for each i, $c_i\in[0,\Pr_x[x\in S_i]]$. Let τ_C be the randomizes threshold function allocating capacity $c_i/\Pr_x[x\in S_i]$ for every group i, as promised from Claim 2.4, and denote by $t_1(c_1),\ldots,t_m(c_m)$ its thresholds. Then

$$\frac{\partial\mathbf{Prop}(\tau_C)}{\partial c_i} = \frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i]} + (\lambda-1)\frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)]}$$

Proof. We have that $\mathbf{Prop}_{\mathcal{D}_p,S}$ is given by

$$\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C) = \sum_{i\in[m]}\Pr_{x\sim\mathcal{D}}[x\in S_i]\log\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i] + (\lambda-1)\log\mathbb{E}_{x\sim\mathcal{D}}[\tau_C(x)p(x)].$$

where τ_C is the randomized threshold function given by Claim 2.4, with thresholds $t_1(c_1),\ldots,t_m(c_m)$ and values $\gamma_1(c_1),\ldots,\gamma_m(c_m)$. The partial derivative is given by:

$$\begin{aligned}
\frac{\partial\mathbf{Prop}_{\mathcal{D}_p,S}(\tau_C)}{\partial c_i} &= \frac{\Pr_{(x,y)\sim\mathcal{D}}[x\in S_i]}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i]}\frac{\partial\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)y|x\in S_i]}{\partial c_i} \\
&\quad+(\lambda-1)\frac{1}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)]}\frac{\partial\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)]}{\partial c_i} \qquad (16)
\end{aligned}$$

We note that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)] = \sum_i\Pr_{(x,y)\sim\mathcal{D}}[x\in S_i]\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i]$. Therefore, we need to find $\frac{\partial}{\partial c_i}\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i]$ to finish the proof.

We denote by Q_i the distribution of $p(x), x\in S_i$. Then let q_i be the quantile function of Q_i and F_i be the CDF of this distribution. In this notation, using Claim A.3:

$$\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x\in S_i] = \Pr_{x\sim\mathcal{D}}[p(x)=t_i(c_i)|x\in S_i]t_i(c_i)\gamma_i(c_i) + \int_{F_i(t_i(c_i))}^{1}q_i(\kappa)d\kappa$$

From Claim 2.4, $t_i(c_i) = q_i(1-w_i)$ for $w_i = \frac{c_i}{\Pr_{x\sim\mathcal{D}}[x\in S_i]}$, and

$$\gamma_i(c_i) = \begin{cases} 0 & \Pr_{u\sim Q_i}[u=t_i]=0, \\ \frac{w_i-1+F_i(t_i))}{\Pr_{u\sim Q_i}[u=t_i]} & \text{o.w.} \end{cases}$$

To calculate the derivative we divide into cases.

1. $\gamma_i(c_i) > 0$: In this case, $\frac{dt_i(c_i)}{dc_i} = 0$, as the quantile function is constant at c_i. The value $\gamma_i(c_i)$ has derivative $1/\Pr[p(x) = t_i, x \in S_i]$, as to increase the total usage of the group by a small amount δ, we need to change γ_i by $\delta/\Pr[p(x) = t_i, x \in S_i]$. That is,

$$\begin{aligned}\frac{\partial \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x \in S_i]}{\partial c_i} &= \Pr_{x\sim\mathcal{D}}[p(x) = t_i(c_i)|x \in S_i]t_i(c_i)\frac{d\gamma_i(c_i)}{dc_i} \\ &= \frac{\Pr_{x\sim\mathcal{D}}[p(x) = t_i(c_i)|x \in S_i]t_i(c_i)}{\Pr[p(x) = t_i, x \in S_i]} \\ &= \frac{t_i(k)}{\Pr_{x\sim\mathcal{D}}[x \in S_i]}.\end{aligned}$$

2. $\gamma_i(c_i) = 0$: the quantile function is not constant at c_i but the first term is zero. In addition, it mean that $F_i(t_i(c_i)) = F_i(q_i(1 - w_i)) = 1 - w_i$. In this case we use the Leibnitz integral rule:

$$\frac{\partial \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x \in S_i]}{\partial c_i} = -q_i(1 - w_i(c_i))\frac{d(1 - w_i(c_i))}{dc_i} = -t_i(c_i)\left(-\frac{1}{\Pr_{x\sim\mathcal{D}}[x \in S_i]}\right).$$

We got that in both cases,

$$\frac{\partial \mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x \in S_i]}{\partial c_i} = \frac{t_i(c_i)}{\Pr_{x\sim\mathcal{D}}[x \in S_i]}.$$

Intuitively, We can see it by noticing that if we increase the total capacity given to S_1 by ϵ, we increase the conditional capacity by $\epsilon/\Pr_{x\sim\mathcal{D}}[x \in S_i]$, and the expected true positive count by this amount the current threshold, as this is the $p(x)$ of the xs at the threshold.

Using this derivative in (16), we get

$$\frac{\partial}{\partial c_i}\mathbf{Prop}(\tau_C) = \frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)|x \in S_i]} + (\lambda - 1)\frac{t_i(c_i)}{\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau_C(x)p(x)]}$$

□

D.2 Max-min fairness and equal opportunity

Theorem D.4. *Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0, 1\}$ with groups S_1, S_2, and suppose $\mathcal{D}$ is $2c$-non-degenerate. Let $p : \mathcal{X} \to [0, 1]$ be an α-calibrated predictor on S_1, S_2, and assume that S_1 has t_0-tail dominance over S_2, with $\Pr[p(x) \geq t_0 \wedge x \in S_i] \geq c$ for $i \in [2]$.*

Let $\tau : [0, 1] \times [2] \to [0, 1]$ be a randomized group threshold function such that $f(x) = \tau(p(x), g(x))$ has capacity c and satisfies max-min fairness on the simulated distribution. Then

$$\mathbb{E}[\tau(p(x), 1) \mid x \in S_1] \;\leq\; \mathbb{E}[\tau(p(x), 2) \mid x \in S_2].$$

Moreover, if τ satisfies max-min fairness up to error δ on $\mathcal{D}$, then

$$\mathbb{E}[\tau(p(x), 1) \mid x \in S_1] \;\leq\; \mathbb{E}[\tau(p(x), 2) \mid x \in S_2] \;+\; \frac{2\alpha + \delta}{t_{\min}},$$

where $t_{\min} \geq t_0$ is the smallest threshold used by τ.

The proof is deferred to Appendix D. The theorem shows that under tail dominance, enforcing max-min fairness shifts allocation toward the less informative group. Next, we strengthen this result by showing that a gap in tail dominance induces a corresponding gap in allocations.

Proof of Theorem D.4. We prove the theorem by applying the technical claim, Claim D.5, appearing below. Let $\mathcal{D}, p$ be such that S_1 has t_0-tail dominance over S_2.

We start with the first part of the theorem, in which we require p to satisfy max-min fairness on the simulated distribution. Let τ be the post-processing, and denote by t_1, t_2 its thresholds on groups

S_1, S_2, respectively. Since τ satisfies max-min fairness in a non-degenerate setting, from Claim 4.3 we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(p(x),1)y \mid x \in S_1] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(p(x),2)y \mid x \in S_2].$$

By the definition of the simulated distribution, this mean that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x) \mid x \in S_1] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x) \mid x \in S_2].$$

The capacity c is such that $\Pr[p(x) \geq t_0 \wedge x \in S_i] \geq c$, which mean that $t_0 \leq t_1, t_2$. Therefore, we can apply Claim D.5 with parameter $\alpha = 0$, and conclude that

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] + \frac{\alpha}{t_{\min}}.$$

For the second part of the theorem, suppose now that τ satisfies max-min fairness on distribution $\mathcal{D}$ that is non-degenerate with error δ, which, using Claim 4.3, implies that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)y \mid x \in S_1] \leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)y \mid x \in S_2] + \delta.$$

From Claim D.1, we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(p(x),1)p(x) \mid x \in S_1] \leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_p}[\tau(p(x),2)p(x) \mid x \in S_2] + 2\alpha + \delta.$$

This is because the distance between the expectation of $\tau(p(x),i)p(x)$ and $\tau(p(x),i)y$ is at most α, the calibration error.

We apply Claim D.5 with parameter $\alpha' = 2\alpha + \delta$, and get

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] + \frac{2\alpha+\delta}{t_{\min}}.$$

□

Claim D.5. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$ where $\mathcal{X}$ is divided into two groups S_1, S_2. Let $p : \mathcal{X} \to [0,1]$ be a predictor on S_1, S_2, such that S_1 is t_0-tail dominant over S_2.

For any randomized group threshold function $\tau : [0,1] \times [2] \to [0,1]$ with thresholds $t_1, t_2 \geq t_0$, if τ on p satisfies

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2] + \alpha \tag{17}$$

then it also satisfies

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] + \frac{\alpha}{t_{\min}},$$

where $t_{\min} = \min\{t_1, t_2\}$.

Proof. Let $\tau : [0,1] \times [2] \to [0,1]$ be a post-processing function satisfying the conditions of the claim. Let t_1, t_2 and γ_1, γ_2 be the thresholds and values of τ on the threshold, for groups S_1, S_2 respectively.

We start by defining the distribution Q_i, which is be the distribution of $p(x)$ when picking $x \sim \mathcal{D}$ conditioning on $x \in S_i$. That is, $\Pr_{u\sim Q_i}[u = r] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) = r|x \in S_i]$. Let F_i be the CDF of this distribution and q_i be the quantile function, see definitions in Appendix A. Using these notations and Claim A.3:

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),i)p(x)|x \in S_i] =& \gamma_i \mathbb{E}[\mathbf{1}(p(x) = t_i)p(x)|x \in S_i] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t_i)p(x)|x \in S_i] \\ =& \gamma_i \mathbb{E}[\mathbf{1}(p(x) = t_i)p(x)|x \in S_i] + \int_{F_i(t)}^{1} q_i(\kappa)d\kappa.\end{aligned}$$

Next, we show that for all $\kappa \in (F_1(t_0), 1]$, $q_1(\kappa) \geq q_2(\kappa)$.

Assume towards a contradiction that there exists $\kappa \in (F_1(t_0), 1]$ such that $q_1(\kappa) < q_2(\kappa)$, and let $t = 0.5(q_1(\kappa) + q_2(\kappa))$. We have that $t_0 < q_1(\kappa) < t < q_2(\kappa)$ (as $\kappa > F_1(t_0)$, see Definition A.2). For this t, by the definition of the quantile function

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) > t | x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[p(x) > q_1(\kappa) | x \in S_1] = 1 - F(q_1(\kappa)) \leq 1 - \kappa,$$
$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) > t | x \in S_2] = 1 - F_2(t) > 1 - \kappa,$$

as if $F_2(t) \geq \kappa$, we would have $q_2(\kappa) \leq t$ contradicting our definition above. The two statements together imply that $\Pr_{(x,y)\sim\mathcal{D}}[p(x) > t | x \in S_1] < \Pr_{(x,y)\sim\mathcal{D}}[p(x) > t | x \in S_2]$, contradicting the tail dominance of S_1.

We use the inequality on the quantile functions to prove the claim, by dividing into a few different cases. The first case is in fact not necessary (covered by the other cases) but as it is the simplest we keep for clarity. The first case is where $p(x)$ is continuous and the high-level idea of the proof appears in it.

1. If $\Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2] = 0$: This is the simplest case, where we have no weight on the thresholds. In this case, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)p(x) | x \in S_i] = \int_{F_i(t_i)}^{1} q_i(\kappa)d\kappa$ and $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i) | x \in S_i] = 1 - F_i(t_i)$. Then, (17) implies

$$\int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa \leq \int_{F_2(t_2)}^{1} q_2(\kappa)d\kappa + \alpha$$

If $F_1(t_1) \geq F_2(t_2)$, we are done. Else, we have

$$\begin{aligned}\int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa &= \int_{F_1(t_1)}^{F_2(t_2)} q_1(\kappa)d\kappa + \int_{F_2(t_2)}^{1} q_1(\kappa)d\kappa \\ &\geq \int_{F_1(t_1)}^{F_2(t_2)} q_1(\kappa)d\kappa + \int_{F_2(t_2)}^{1} q_2(\kappa)d\kappa,\end{aligned}$$

where we use the fact that $q_1(\kappa) \geq q_2(\kappa)$ for all κ in the range. From both inequalities we conclude that $\int_{F_1(t_1)}^{F_2(t_2)} q_1(\kappa)d\kappa \leq \alpha$, implying that $(F_2(t_2) - F_1(t_1))t_1 \leq \alpha$, where we use Claim A.4 to lower bound the integral. Therefore, we have that $F_1(t_1) \geq F_2(t_2) - \alpha/t_1$.

2. If $t_1 < t_2$: set $t = 0.5(t_1 + t_2)$, then $t_0 \leq t_1 < t < t_2$. We express the true positive count using t:

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 1)p(x) | x \in S_1] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1 t_1 | x \in S_1] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t_1)p(x) | x \in S_1] \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1 t_1 | x \in S_1] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(t_1 < p(x) \leq t)p(x) | x \in S_1] \\ &\quad + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t)p(x) | x \in S_1] \\ \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 2)p(x) | x \in S_2] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 t_2 | x \in S_2] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t_2)p(x) | x \in S_2] \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t)p(x) | x \in S_2] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(t_2 < p(x) < t)p(x) | x \in S_2] \\ &\quad - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)t_2(1 - \gamma_2) | x \in S_2]\end{aligned}$$

Since $q_1(\kappa) \geq q_2(\kappa)$ for all κ in range, we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t)p(x) | x \in S_2] = \int_{F_2(t)}^{1} q_2(\kappa)d\kappa \leq \int_{F_1(t)}^{1} q_1(\kappa)d\kappa = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) > t)p(x) | x \in S_1].$$

Therefore, in order to satisfy (17) we must have:

$$\begin{aligned}\alpha \geq & \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(t_1 < p(x) \leq t)p(x)|x \in S_1] + \gamma_1 \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)t_1|x \in S_1] \\ & + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(t_2 < p(x) < t)p(x)|x \in S_2] + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)t_2(1-\gamma_2)|x \in S_2] \\ \geq & \int_{F_1(t_1)}^{F_1(t)} q_1(\kappa)d\kappa + \int_{F_2(t)}^{F_2(t_2)} q_2(\kappa)d\kappa + t_1 \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1] \\ & - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2|x \in S_2]t_1\end{aligned}$$

We lower bound the integrals by $(F_1(t) - F_1(t_1))t_1 + (F_2(t_2) - F_2(t))t_2$ using Claim A.4. Since we have that $t_1 < t_2$ and $F_2(t) \geq F_1(t)$, we can lower bound the integrals by $(F_2(t_2) - F_1(t_1))t_1$.

Therefore, we conclude that

$$\begin{aligned}\alpha \geq & (F_2(t_2) - F_1(t_1))t_1 + t_1\gamma_1 \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1|x \in S_1] - \gamma_2 \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2|x \in S_2]t_1 \\ \geq & \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1]t_1 - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2]t_1,\end{aligned}$$

finishing the claim for this case.

3. If $t_1 \geq t_2, F_1(t_1) \leq F_2(t_2)$: Denote $\delta = F_2(t_2) - F_1(t_1)$.

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1 t_1|x \in S_1] + \int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1 t_1|x \in S_1] + \int_{F_1(t_1)}^{F_2(t_2)} q_1(\kappa)d\kappa + \int_{F_2(t_2)}^{1} q_1(\kappa)d\kappa \\ \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_2] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2] + \int_{F_2(t_2)}^{1} q_2(\kappa)d\kappa\end{aligned}$$

We know that $q_1(\kappa) \geq q_2(\kappa)$ for all κ in our range. In order to satisfy (17):

$$\begin{aligned}\alpha + \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2|x \in S_2]t_2 &\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1]t_1 + \int_{F_1(t_1)}^{F_2(t_2)} q_1(\kappa)d\kappa \\ &\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1]t_1 + \delta t_1,\end{aligned}$$

where the last inequality is from Claim A.4. Since $t_1 \geq t_2$, which means

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2|x \in S_2] \geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1] + \delta - \frac{\alpha}{t_2}.$$

which is enough to finish the claim in this case.

4. If $t_1 \geq t_2, F_1(t_1) > F_2(t_2)$: Denote $\Delta = F_1(t_1) - F_2(t_2)$.

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1 y|x \in S_1] + \int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa \\ \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_2] &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2] + \int_{F_2(t_2)}^{1} q_2(\kappa)d\kappa \\ &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2] + \int_{F_2(t_2)}^{F_1(t_1)} q_2(\kappa)d\kappa + \int_{F_1(t_1)}^{1} q_2(\kappa)d\kappa.\end{aligned}$$

To satisfy (17) we must have:

$$\begin{aligned}\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1]t_1 &\leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2]t_2 + \int_{F_2(t_2)}^{F_1(t_1)} q_2(\kappa)d\kappa + \alpha \\ &\leq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2]t_2 + \Delta t_2 + \alpha\end{aligned}$$

Since $t_1 \geq t_2$ we have that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_1)\gamma_1|x \in S_1] \leq \mathbb{E}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(p(x) = t_2)\gamma_2 y|x \in S_2] + \Delta + \alpha/t_1$, which finishes the proof. □

Proof of Theorem 5.3. Let $\mathcal{D}, p$ be as in the theorem statement. We prove the theorem using the technical claim, Claim appearing below.

The predictor p is such that S_1 has (t_0, t_M) tail dominance over S_2, with gap η. That is, it satisfies (19) and (20) in the requirements of . In addition, from the theorem statement we have that p and c satisfies (22) (18) and (21) and that $c \geq 6\eta$ and we are in a non-degenerate setting. Therefore, we have that p satisfies all the requirements of Claim D.6.

We start with the first part of the theorem statement. Assume that τ satisfies max-min fairness in a non-degenerate setting with respect to the simulated distribution, i.e. using Claim 4.3 (where we are on the simulated distribution so $y \sim \text{Ber}(p(x))$),

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2].$$

Then from Claim D.6 with $\alpha = 0$ we get that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \geq \frac{\beta\eta}{t_2},$$

where $t_2 \leq t_1$.

For the second part of the theorem, we assume that τ satisfies max-min fairness up to an error δ on the distribution $\mathcal{D}$ that is non-degenerate. Using Claim 4.3 we have that

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)y|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)y|x \in S_2] + \delta.$$

Since p is α-calibrated, we have that

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2] + 2\alpha + \delta.$$

Therefore, we can apply Claim D.6 with error parameter $2\alpha + \delta$ and we get

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \geq \frac{\beta\eta - 2\alpha - \delta}{t_2},$$

where $t_2 \leq t_1$. □

Claim D.6. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$ where $\mathcal{X}$ is divided into two groups S_1, S_2. Let $p : \mathcal{X} \to [0,1]$ be a predictor on S_1, S_2. Let $t_0, t_M, \eta, \beta, \alpha, c \in (0,1)$ such that $\alpha < \beta\eta, c \geq 6\eta$ be constants such that for $i \in [2]$:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) \geq t_M|x \in S_2] \leq \eta \tag{18}$$

$$\forall r \in [t_0, t_M] \quad \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) > r|x \in S_1] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) > r|x \in S_2] \geq \eta \tag{19}$$

$$\forall r \geq t_M \quad \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) > r|x \in S_1] \geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) > r|x \in S_2] \tag{20}$$

$$\forall r \geq t_0 \quad \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) \in [r, r+\beta]|x \in S_i] < \eta \tag{21}$$

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) \geq t_0 \wedge x \in S_i] \geq c \tag{22}$$

Let $\tau : \mathcal{X} \times [2] \to [0,1]$ be a randomized threshold function with thresholds t_1, t_2 that satisfies

$$\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] \leq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2] + \alpha. \tag{23}$$

Then τ also satisfies

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] \geq \frac{\beta\eta - \alpha}{t_2}, \tag{24}$$

where $t_2 \leq t_1$ is the threshold of τ on group 2.

Proof. The proof of this claim follows the same high-level approach of Claim D.5, but preserves the gap using the additional assumption on the density of the distribution defined by $p(x), x \sim \mathcal{D}$. Let $t_1, t_2, \gamma_1, \gamma_2$ be the thresholds and values of τ.

Following the approach of the previous proof, let F_i be the comulative distribution function of $p(x)$ for $x \sim \mathcal{D}$ conditioning on $x \in S_i$, and let q_i be the quantile function of this distribution. By the definition of τ and Claim A.4 for $i \in \{1, 2\}$,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)p(x)|x \in S_i] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_i|x \in S_i]\gamma_i \cdot t_i + \int_{F_i(t_i)}^{1} q_i(\kappa)d\kappa.$$

Let $K_0 = F_1(t_0)$. We start by showing that for all $\kappa \in (K_0, 1-2\eta)$ we have that $q_1(\kappa) \geq q_2(\kappa) + \beta$.

Assume towards a contradiction that there is some $\kappa \in (K_0, 1-2\eta)$ such that $q_1(\kappa) < q_2(\kappa) + \beta$, and set $t = 0.5(q_1(\kappa) + q_2(\kappa) + \beta)$. Then $t > q_1(\kappa) \geq t_0$. We show that $t < t_M$: if $t \geq t_M$ then $F_2(t) \geq F_2(t_M) \geq 1 - \eta$. But, we also have that $F_2(t) \leq F_2(t - \beta) + \Pr_{(x,y)\sim\mathcal{D}}[p(x) \in [t - \beta, t]] < 1 - 2\eta + \eta$, where we use (21) and the fact that $t - \beta < q_2(1 - 2\eta)$. Therefore, we can use (19) on t:

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) > t|x \in S_1] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) > q_1(\kappa)|x \in S_1] + \Pr_{(x,y)\sim\mathcal{D}}[p(x) \in (t, q_1(\kappa)]|x \in S_1] < 1 - \kappa + \eta,$$
$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) > t|x \in S_2] = 1 - F_2(t) \geq 1 - \kappa.$$

Where in the top inequality we use (21). In the second inequality we use the fact that $t > q_2(\kappa)$. The two inequalities contradict (19), and therefore we conclude that for all $\kappa \in (K_0, 1-2\eta)$, we have that $q_1(\kappa) \geq q_2(\kappa) + \beta$.

We show next that $t_1 \geq t_2$. Assume towards a contradiction that $t_1 < t_2$. From the conditions of the claim, we have that $t_1 \geq t_0$, therefore from the monotonicity,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 1)|x \in S_1] \geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 2)|x \in S_2].$$

Since our capacity is $c > 6\eta$, then $\mathbb{E}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 1)|x \in S_1] > 6\eta$. This mean that $F_1(t_1) \geq \mathbb{E}_{(x,y)\sim\mathcal{D}}[p(x) \geq t_1|x \in S_1] - \mathbb{E}_{(x,y)\sim\mathcal{D}}[p(x) = t_1|x \in S_1] \geq 5\eta$. Where we use the smoothness condition to show that the capacity at a single point is bounded. From the quantile inequality, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 1)p(x)|x \in S_1] &\geq \int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa \\
&\geq \int_{F_1(t_1)}^{1-2\eta} q_1(\kappa)d\kappa + \int_{1-2\eta}^{1} q_1(\kappa)d\kappa \\
&\geq \int_{F_1(t_1)}^{1-2\eta} q_2(\kappa) + \beta d\kappa + \int_{1-2\eta)}^{1} q_2(\kappa)d\kappa \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), 2)p(x)|x \in S_2] + \beta\eta.
\end{aligned}$$

This is a contradiction to the claim as $\alpha < \beta\eta$. Therefore, in the rest of the proof, we have that $t_1 \geq t_2$.

We prove that $\min\{F_1(t_1), F_2(t_2)\} \leq 1 - 4\eta$. Let $i_{\min}$ be the indicator to the index of the minimal $F_i(t_i)$. For every i,

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), i)|x \in S_i] &\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) > t_i|x \in S_i] - \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[p(x) = t_i|x \in S_i] \\
&\geq 1 - F_i(t_i) - \eta
\end{aligned}$$

Using the capacity bound,

$$\begin{aligned}
c &= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x), g(x))] \\
&\geq \sum_{i\in[2]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i](1 - F_i(t_i) - \eta) \\
&\geq 1 - F_{i_{\min}}(t_{i_{\min}}) - \eta
\end{aligned}$$

We conclude that $1 - F_{i_{\min}}(t_{i_{\min}}) \geq c - \eta \geq 4\eta$.

We divide into cases:

1. If $F_1(t_1) \leq F_2(t_2)$, we denote $F_2(t_2) = F_1(t_1) + \delta$, and note that in this case, $1 - F_1(t_1) \geq 4\eta$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] &= \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 \cdot t_1 + \int_{F_1(t_1)}^{1-2\eta} q_1(\kappa)d\kappa + \int_{1-2\eta}^{1} q_1(\kappa)d\kappa \\
&\geq \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 \cdot t_1 + \int_{F_1(t_1)}^{1-2\eta} q_2(\kappa)d\kappa \\
&\quad + \beta(1 - F_1(t_1) - 2\eta) + \int_{1-2\eta}^{1} q_2(\kappa)d\kappa. \\
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2] &= \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 \cdot t_2 + \int_{F_2(t_2)}^{1-2\eta} q_2(\kappa)d\kappa + \int_{1-2\eta}^{1} q_2(\kappa)d\kappa
\end{aligned}$$

   Requiring (23) implies that

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 \geq \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 + \delta + \frac{\beta(1 - F_1(t_1) - 2\eta) - \alpha}{t_2}.$$

   We note that $1 - F_1(t_1) \geq 4\eta$ and that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 + (1 - F_2(t_2))$$

   and finish the proof for this case.

2. If $F_1(t_1) > F_2(t_2), t_1 \geq t_2$, we denote $F_1(t_1) = F_2(t_2) + \Delta$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)p(x)|x \in S_1] &= \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 t_1 + \int_{F_1(t_1)}^{1-2\eta} q_1(\kappa)d\kappa + \int_{1-2\eta}^{1} q_1(\kappa)d\kappa \\
&\geq \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 \cdot t_1 + \int_{F_1(t_1)}^{1-2\eta} q_2(\kappa)d\kappa \\
&\quad + \beta(1 - 2\eta - F_1(t_1)) + \int_{1-2\eta}^{1} q_2(\kappa)d\kappa. \\
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)p(x)|x \in S_2] &= \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 t_2 + \int_{F_2(t_2)}^{F_1(t_1)} q_2(\kappa)d\kappa \\
&\quad + \int_{F_1(t_1)}^{1-2\eta} q_2(\kappa)d\kappa + \int_{1-2\eta}^{1} q_2(\kappa)d\kappa \\
&\leq \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 \cdot t_2 + t_1\Delta \\
&\quad + \int_{F_2(t_2)}^{1-2\eta} q_2(\kappa)d\kappa + \int_{1-2\eta}^{1} q_2(\kappa)d\kappa
\end{aligned}$$

   Requiring (23) implies that

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 \geq \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_1 | x \in S_1]\gamma_1 - \Delta + \frac{\beta(1 - 2\eta - F_2(t_2)) - \alpha}{t_2}.$$

   Substituting this in the expectation, we have that

$$\begin{aligned}
\Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),2)|x \in S_2] &= \Pr_{(x,y)\sim\mathcal{D}}[p(x) = t_2 | x \in S_2]\gamma_2 + \Pr_{(x,y)\sim\mathcal{D}}[p(x) > t_2 | x \in S_2] \\
&\geq \Pr_{(x,y)\sim\mathcal{D}}[\tau(p(x),1)|x \in S_1] + \frac{\beta\eta - \alpha}{t_2},
\end{aligned}$$

   finishing the proof for this case.

□

Claim D.7. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0,1\}$, and assume that $\mathcal{X}$ is divided into m disjoint groups $S_1, \ldots, S_m$. Let p be a predictor on the groups $S_1, \ldots, S_m$. Then, the randomized threshold function that has a single threshold to all groups and has capacity c maximizes the true positive count.

Proof. Let τ' be the randomized threshold function that applied a single threshold to all groups and has capacity c. From Claim 2.4, there must be such function τ'. Denote by t' the threshold of τ'.

Let τ be any group post-processing of p that has capacity c. Denote by S the set of x's for which $\tau(p(x), g(x)) > \tau'(p(x), g(x))$ and by S' those for which $\tau(p(x), g(x)) < \tau'(p(x), g(x))$.

Since τ, τ' have equal capacity, by the definition of S, S' we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(\tau(p(x), g(x)) - \tau'(p(x), g(x)))\mathbf{1}(x \in S)] = \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[(\tau'(p(x), g(x)) - \tau(p(x), g(x)))\mathbf{1}(x \in S')].$$

By definition, for all x such $p(x) > t'$ we have $\tau'(p(x), g(x) = 1$, therefore for all $x \in S$ we have that $p(x) \leq t$. Similarly, for all x such that $p(x) < t'$ we have that $\tau'(p(x), g(x)) = 0$, implying that for all $x \in S'$, $p(x) \geq t'$. We use it to show that τ' has maximal true positive count for the simulated distribution:

$$\begin{aligned}
\mathbb{E}[(\tau'(p(x), g(x)) - \tau(p(x), g(x)))p(x)] =& \mathbb{E}[(\tau'(p(x), g(x)) - \tau(p(x), g(x)))p(x)\mathbf{1}(x \in S')] \\
&- \mathbb{E}[(\tau(p(x), g(x)) - \tau'(p(x), g(x)))p(x)\mathbf{1}(x \in S)] \\
\geq& t\,\mathbb{E}[(\tau'(p(x), g(x)) - \tau(p(x), g(x)))\mathbf{1}(x \in S')] \\
&- t\,\mathbb{E}[(\tau(p(x), g(x)) - \tau'(p(x), g(x)))\mathbf{1}(x \in S)] \\
\geq& 0.
\end{aligned}$$

Where all expectations are over $x \sim \mathcal{D}$. □

Proof of Claim 5.4. Let $\mathcal{D}, p$ be as in the conditions of the claim. Let τ be the randomized group threshold function satisfying max-min fairness, and let t_1, t_2 be the thresholds for groups S_1, S_2 respectively. Let τ' be the randomized threshold function that uses a single threshold for all groups and has capacity c. From Claim D.7, τ' maximizes the true positive count under the simulated distribution. From Claim D.1, it maximizes the true positive count under $\mathcal{D}$ up to an additive error of 2α.

Since S_1 is tail-dominant over S_2 and from our capacity $t' \in [t_0, t_M]$, we have that $F_1(t') \leq F_2(t') - \eta$. From Theorem 5.3, $F_1(t_1) \geq F_2(t_2) + (\eta\beta - 2\alpha - \delta)/t_2$.

We know that τ' and τ have the same capacity c. Since the probability distribution of p is continuous, we get that

$$\begin{aligned}
c =& \Pr_{x\sim\mathcal{D}}[x \in S_1](1 - F_1(t_1)) + \Pr_{x\sim\mathcal{D}}[x \in S_2](1 - F_2(t_2)) \\
F_1(t_1) \geq& 1 - c + \frac{\eta\beta - 2\alpha - \delta}{t_2} \Pr_{x\sim\mathcal{D}}[x \in S_2] \\
F_2(t_2) \leq& 1 - c - \frac{\eta\beta - 2\alpha - \delta}{t_2} \Pr_{x\sim\mathcal{D}}[x \in S_1] \\
c =& \Pr_{x\sim\mathcal{D}}[x \in S_1](1 - F_1(t')) + \Pr_{x\sim\mathcal{D}}[x \in S_2](1 - F_2(t')) \\
F_1(t') \leq& 1 - c - \eta \Pr_{x\sim\mathcal{D}}[x \in S_2] \\
F_2(t') \geq& 1 - c + \eta \Pr_{x\sim\mathcal{D}}[x \in S_1].
\end{aligned}$$

Since our groups are equal-sized, we have that $F_1(t_1) - F_1(t') \geq \left(\frac{\eta\beta - 2\alpha - \delta}{t_2} + \eta\right)\frac{1}{2}$, and $F_1(t_1) - F_1(t') = F_2(t') - F_2(t_2)$.

Since the distribution of $p(x), x \sim \mathcal{D}$ is continuous for $x \in S_1, S_2$, we have that

$$\Pr_{(x,y)\sim\mathcal{D}}[p(x) \in (t', t_1), x \in S_1] = \Pr_{(x,y)\sim\mathcal{D}}[p(x) \in (t_2, t'), x \in S_2],$$

Using Claim A.3,

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x),g(x))p(x)] &= \Pr_{x\sim\mathcal{D}}[x\in S_1]\int_{F_1(t')}^{1} q_1(\kappa)d\kappa + \Pr_{x\sim\mathcal{D}}[x\in S_2]\int_{F_2(t')}^{1} q_2(\kappa)d\kappa \\
&= \Pr_{x\sim\mathcal{D}}[x\in S_1]\int_{F_1(t')}^{F_1(t_1)} q_1(\kappa)d\kappa + \Pr_{x\sim\mathcal{D}}[x\in S_1]\int_{F_1(t_1)}^{1} q_1(\kappa)d\kappa \\
&\quad + \Pr_{x\sim\mathcal{D}}[x\in S_2]\int_{F_2(t_2)}^{1} q_2(\kappa)d\kappa - \Pr_{x\sim\mathcal{D}}[x\in S_2]\int_{F_2(t_2)}^{F_2(t')} q_2(\kappa)d\kappa \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] + \Pr_{x\sim\mathcal{D}}[x\in S_1]\int_{F_1(t')}^{F_1(t_1)} q_1(\kappa)d\kappa \\
&\quad - \Pr_{x\sim\mathcal{D}}[x\in S_2]\int_{F_2(t_2)}^{F_2(t')} q_2(\kappa)d\kappa
\end{aligned}$$

We bound the integrals. From the Lipchitz condition, at most $\eta/4$ of the weight can be in a range $[r, r+\beta]$. Therefore, we can lower bound the probability as follows:

$$\begin{aligned}
\mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\mathbf{1}(p(x)\in(t',t_1))p(x)|x\in S_1] &\geq \frac{\eta}{4}t' + \left(\Pr_{x\sim\mathcal{D}}[p(x)\in(t',t_1)|x\in S_1] - \frac{\eta}{4}\right)(t'+\beta) \\
&\geq \Pr_{x\sim\mathcal{D}}[p(x)\in(t',t_1)|x\in S_1](t'+\beta) - \frac{\eta\beta}{4}.
\end{aligned}$$

For the second integral,

$$\mathop{\mathbb{E}}_{x\sim\mathcal{D}}[\mathbf{1}(p(x)\in(t_2,t'))p(x)|x\in S_2] \leq t' \Pr_{x\sim\mathcal{D}}[p(x)\in(t_2,t')|x\in S_2]$$

Combining it together we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x),g(x))p(x)] &\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] + \Pr_{x\sim\mathcal{D}}[p(x)\in(t',t_1), x\in S_1](t'+\beta) \\
&\quad - \Pr_{x\sim\mathcal{D}}[x\in S_1]\frac{\beta\eta}{4} - t' \Pr_{x\sim\mathcal{D}}[p(x)\in(t_2,t'), x\in S_2] \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] \\
&\quad + \Pr_{x\sim\mathcal{D}}[x\in S_1]\beta\left(\Pr_{x\sim\mathcal{D}}[p(x)\in(t',t_1)|x\in S_1] - \frac{\eta}{4}\right) \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] + \frac{\eta\beta}{8} + \frac{\beta(\eta\beta-2\alpha-\delta)}{4t_2}.
\end{aligned}$$

Where we use the fact that the groups are equal sized, and $F_1(t_1) - F_1(t') \geq \left(\frac{\eta\beta-2\alpha-\delta}{t_2} + \eta\right)\frac{1}{2}$ from above.

Using Claim D.1, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))y] &\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x),g(x))y] - 2\alpha \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau'(p(x),g(x))p(x)] - 3\alpha \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))p(x)] + \frac{\eta\beta}{8} + \frac{\beta(\eta\beta-4\alpha-\delta)}{4t_2} - 3\alpha \\
&\geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[\tau(p(x),g(x))y] + \frac{\eta\beta}{8} + \frac{\beta(\eta\beta-2\alpha-\delta)}{4t_2} - 4\alpha
\end{aligned}$$

□

E Proofs of Section 6

Proof of Claim 6.2. For every $i \in [m]$, let $f_i = \arg\max_{f'\in\mathcal{F}_c} \mathbb{E}_{(x,y)\sim\mathcal{D}}[f'(x)y]$. Let f be the function defined by $f^*(x) = f_{g(x)}(x)$. Since $\mathcal{F}$ is closed under group choice post-processing, we

have that $f^* \in \mathcal{F}$. We note that since each f_i can allocate all of the resources only to S_i, the capacity of f^* can be mc. By definition, $\mathbb{E}[f^*(x)y] = \sum_{i\in[m]} \Pr_{(x,y)\sim\mathcal{D}}[x \in S_i]\mathcal{A}(\mathcal{F}, c, S_i)$.

Since h has capacity c over all groups and f^* picks for every group the function in $\mathcal{F}$ with capacity c that maximizes the true positive count, we have that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f^*(x)y] \geq \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y].$$

We prove next that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y] \geq \mathbb{E}_{(x,y)\sim\mathcal{D}}[f^*(x)y]/m$, which finished the proof.

We start by explicitly writing the true positive count of f, the function satisfying achievable equal opportunity:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y] = \sum_j \Pr_{(x,y)\sim\mathcal{D}}[x \in S_j] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i].$$

Let $i \in [m]$ be an arbitrary index, then from the achievable equal opportunity requirement,

$$\begin{aligned}
\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y] =& \sum_j \Pr_{(x,y)\sim\mathcal{D}}[x \in S_j] \frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i]\mathcal{A}(\mathcal{F}, c, S_j)}{\mathcal{A}(\mathcal{F}, c, S_i)} \\
=& \frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i]}{\mathcal{A}(\mathcal{F}, c, S_i)} \sum_j \Pr_{(x,y)\sim\mathcal{D}}[x \in S_j]\mathcal{A}(\mathcal{F}, c, S_j) \\
=& \frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i]}{\mathcal{A}(p, c, S_i)} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f^*(x)y].
\end{aligned}$$

The above equality holds for every i, and specifically let $i^* \in [m]$ be such that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)|x \in S_{i^*}] \geq c$ (by averaging, we must have at least one such i^* as the capacity of f is c). For i^*,

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_{i^*}] \geq \mathcal{A}(\mathcal{F}, c, S_{i^*})/m.$$

This is because $\mathcal{F}$ is closed under group choice post-processing, and therefore we could have set $f(x) = 1/m f_{i^*}(x)$ for all $x \in S_{i^*}$ and 0 for all other groups. It is possible, of course, that there is a better choice for f on S_{i^*}, and in this case the inequality will be strict.

Combining everything together, we get that

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y] \geq \frac{1}{m} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f^*(x)y] \geq \frac{1}{m} \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y].$$

□

We show next that the above bound is tight, and we can find m groups such that requiring achievable equal opportunity reduces the expected true positive count by a factor approaching m.

Example E.1. Let $\mathcal{D}$ be a distribution over $\mathcal{X} \times \{0, 1\}$, and assume that $\mathcal{X}$ is divided into m equal-sized groups. Let $\delta < 1/m$ be a small constant, and let $k < 1/(m\delta c)$. Let $p : \mathcal{X} \to [0, 1]$, such that on all $x \in S_1$, $p(x) = k\delta$ and for all $i > 1$ we have $p(x) = \delta$ for all $x \in S_i$. This mean that for all $c \leq 1/m$, $\mathcal{A}(p, c, S_1) = k\delta mc$ and for all $i > 1$,$\mathcal{A}(p, c, S_i) = \delta mc$.

The function $h : \mathcal{X} \to [0, 1]$ that has the maximal true positive count has $h(x) = 0$ to all $x \notin S_1$ and $h(x) = cm$ for all $x \in S_1$. It satisfies:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y] = \Pr_{(x,y)\sim\mathcal{D}}[x \in S_1] \mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[h(x)y|x \in S_1] = \frac{1}{m} ck\delta m = ck\delta.$$

Let f be the function maximizing the true positive count while satisfying the achievable equal opportunity constraint. For such f, by definition,

$$\frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_1]}{\delta kmc} = \frac{\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i]}{\delta mc}.$$

	Max		Max-Min		Prop (λ=1)		Prop (λ=10)	
Capacity	Frac	Util	Frac	Util	Frac	Util	Frac	Util
0.005	1.000	0.0021	0.330	0.0016	0.499	0.0017	0.800	0.0020
0.010	1.000	0.0040	0.328	0.0029	0.499	0.0033	0.793	0.0037
0.020	0.999	0.0072	0.326	0.0054	0.499	0.0060	0.786	0.0069
0.040	0.995	0.0130	0.323	0.0099	0.498	0.0110	0.773	0.0124
0.100	0.956	0.0269	0.317	0.0214	0.497	0.0238	0.744	0.0261
0.200	0.855	0.0446	0.307	0.0368	0.495	0.0411	0.703	0.0439

Table 1: Tradeoff between allocation (fraction of the capacity assigned to group 1) and utility across objectives.

That is, we must have $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_1] = k\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i]$. For our function p, this implies that $f(x) = c$ for all x, as then we have $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)] = c$, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_1] = k\delta c$ and for all $i > 1$, $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y|x \in S_i] = \delta c$ as required by the equal opportunity condition.

The true positive count of f is given by:

$$\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}}[f(x)y] = \frac{1}{m}k\delta c + \frac{m-1}{m}\delta c \leq \delta c\left(\frac{k}{m}+1\right).$$

When k is large, we get that $\mathbb{E}_{(x,y)\sim\mathcal{D}}[f(x)y] \approx \frac{1}{m}\,\mathbb{E}_{(x,y)\sim\mathcal{D}}[h(x)y]$, as required.

F Simulation

We describe the simulations appearing in the figures, and present more simulations.

Figure 1a: This figure is a histogram of a distribution of a predictor $p(x)$. We generate the distribution of $p(x)$ on group S_1 as a Gaussian random variable with mean 0.02 and variance 0.02, and the distribution of $p(x)$ on group S_2 is a Gaussian random variable with mean 0.01 and variance 0.005. Both distributions are clipped to $[0, 1]$ (since they represent the output of a predictor), so the actual mean after clipping is different.

The red vertical line is a cutoff of the top 1% of the population. Nearly 100% of the individuals above the cutoff are from S_1.

Figure 1b: For every fairness definition and every capacity c in $0.005, 0.01, 0.02, 0.04, 0.1, 0.2$, the graph has a point for the ratio between the total utility under the fairness requirement and the maximal utility for this capacity, under the same distribution as Figure 1a. For every fairness definition, if f_c is the allocation of f satisfying the fairness definition and having capacity c, and h_c is the maximal utility allocation with capacity c, then the ratio is:

$$\frac{\mathbb{E}[f_c(x)y]}{\mathbb{E}[h_c(x)y]}$$

for every capacity c and every fairness requirement in the range.

The fraction of resources given to S_1 comparing to S_2 for each of the fairness definitions and capacities is given in Table 1. We can see that max-min fairness allocates only about third of the capacity to the dominant group S_1, the maximal utility allocates all of it or the large majority (depending on the total capacity), proportional fairness about half, and the regularization proportional fairness about $70 - 80\%$. Similarly, the utility grows with the capacity to the dominant group.

Figure 2: This graph shows a comparison between the theoretical bounds of Theorem 5.1 and the bounds on the distribution of Figure 1a with two equal-sized groups. We use the capacity of 1%, and measure the ratio between the λ-proportional fairness allocation with and the maximal utility, for λ in $1, 2, 4, 8, 16, 32$. The theoretical bound with $\mu = 0.5$ for two equal sized groups is in the dashed lines, and it is a lower bound on the actual bound we have in practice. This lower bound is an upper bound on the price of fairness, as it is a lower bound for the ratio between the fair utility and the maximal achievable utility.

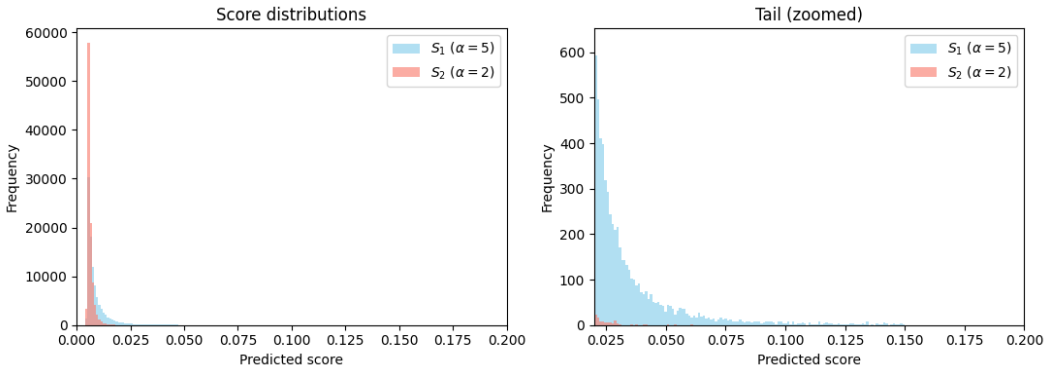


Figure 3: Score distributions for two groups generated from Pareto distributions with scale 0.005 and shape parameters $\alpha_1 = 5$ and $\alpha_2 = 2$. Left: the full distributions. Right: a zoomed view of the upper tails.

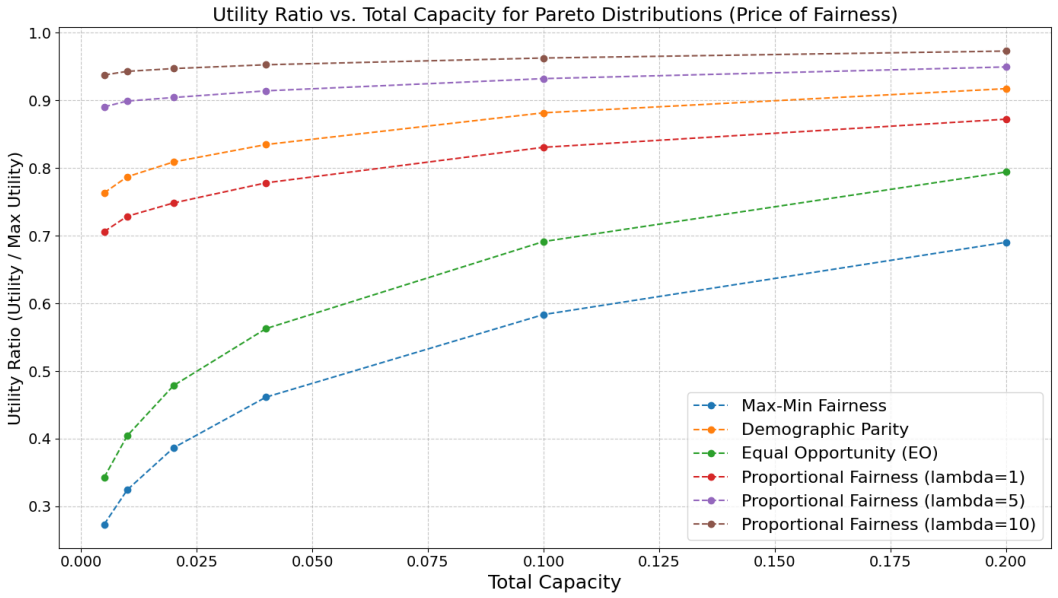


Figure 4: Price of fairness of demographic parity, max-min fairness, equal opportunity and proportional fairness with varying values of λ for the heavy-tail distribution described above. The graph contains the utility ratio for the given distributions.

Heavy-tail distributions To verify the robustness of the results, we preformed simulations also on long-tail distribution, in addition to the Gaussian distribution appearing in the body of the paper. The chosen long-tail distribution is Pareto distribution, with scale 0.005 and group S_1 with index 5 and group S_2 with index 2. Attached the histogram of the two distributions, in figure 3.

The price of fairness of the different fairness definitions is in Figure 4. The graph presents the price of fairness multiplicatively, that is, ratio between the utility of satisfying the fairness definition divided by the maximal achievable utilities, for different capacity values. In Figure 5 we compare the price of fairness of the different regularization parameters of the proportional fairness definitions. The heavy-tail distributions have a relatively higher price of fairness comparing to the Gaussian distribution, since the heavy tail increases the gaps between the risk scores of the chosen individuals in the two groups. Yet, the strictness of the fairness definition is preserved at large. The only difference is that for these distributions, proportional fairness with $\lambda = 1$ is more strict than demographic parity.

In addition, in Figure 5 we compared the price of fairness of proportional fairness to the theoretical bound for the long-tail distribution. This is the same comparison of the theoretical bounds as in Figure

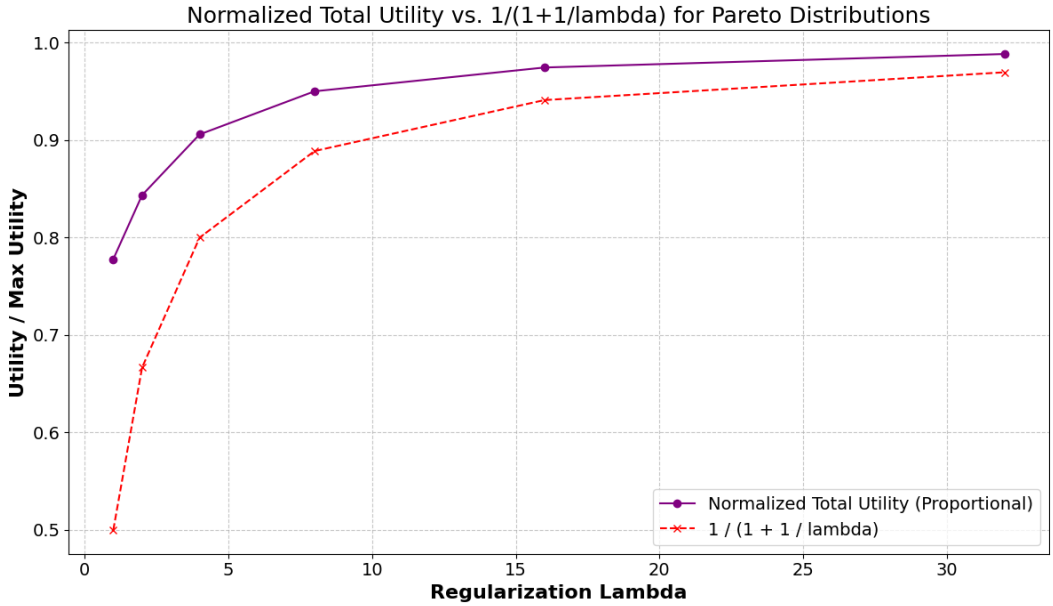


Figure 5: The price of fairness for proportional fairness with different values of λ and the theoretial lower bound, for capacity 0.01 and the Pareto distribution above.

2, on a heavy-tail distribution. The simulations here has a higher price of fairness comparing to the Gaussian distribution, but still lower than the theoretical bounds. This is expected as the theoretical bound holds for every distribution.

Shifted Gaussian distribution In addition to adding heavy-tail distributions, we also simulate shifted Gaussian. These are meant to model a setting where we have the same "shape" of distribution on the two groups, but one has systematically higher risk. In Figure 6 we have the chosen shifted distribution, where in S_1 we have Gaussian distribution with mean 0.07 and standard deviation 0.05 and in S_2 mean 0.02 and standard deviation 0.05.

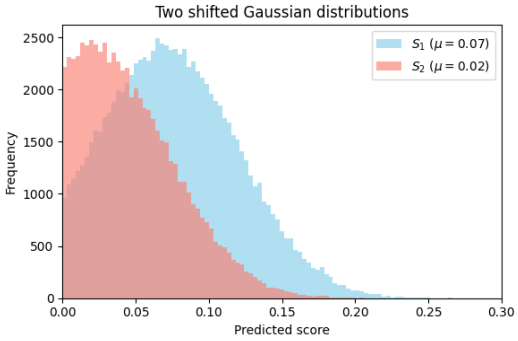


Figure 6: The histogram of two Gaussian distribution with equal variance and different mean, the 0 bin is clipped out.

We plot the same graphs as in the paper, for the multiplicative price of fairness of the different distributions, for different capacities in Figure 7. We can see that the shifted Gaussian is slightly different, in the sense that max-min fairness is still the most strict, but Equal opportunity is much less. This is because in the current distribution, the group S_2 is not more concentrated around the mean, as in the previous distributions, but has the same shape but shifted. This translates to a significant gap in the true positive $y = 1$ between groups S_1, S_2. While max-min fairness still requires the same true positive count $f(x)y$, Equal opportunity normalize it by the total number of $y = 1$ in the group. Therefore, in this case where we have very few $y = 1$ in group S_2 comparing to S_1, it is less strict.

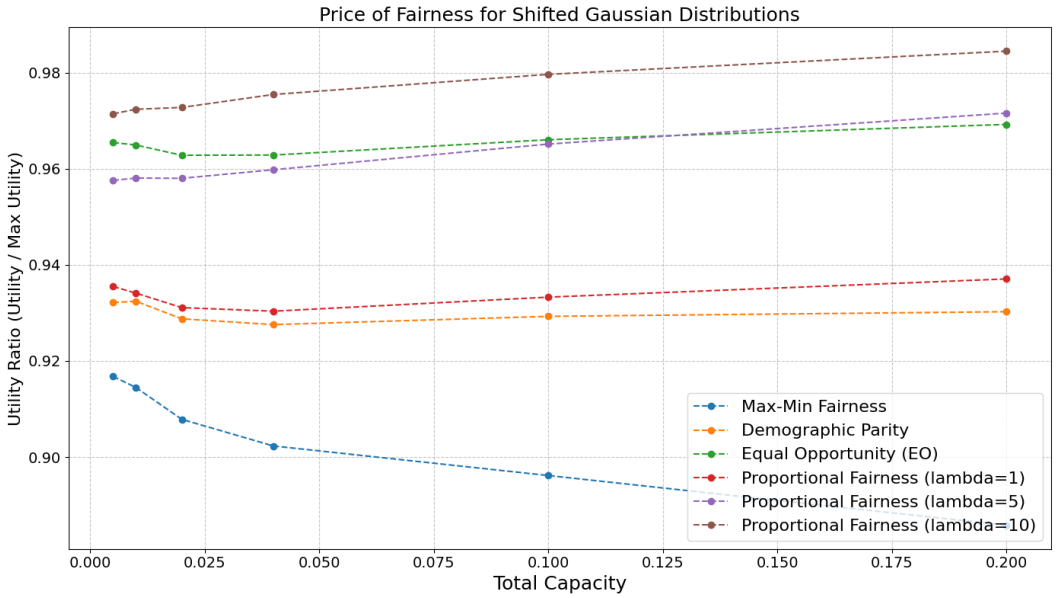


Figure 7: Price of fairness of demographic parity, max-min fairness, equal opportunity and proportional fairness with varying values of λ for the shifted Gaussian distribution described above. The graph contains the utility ratio for the given distributions.

Similarly, we show the theoretical bound on the price of fairness of proportional fairness with different values of λ, in Figure 8. In this distribution the price of fairness of all fairness requirements is rather small, resulting in a larger gap between the lower bound price-of-fairness, that is worse case, and the simulated price of fairness for this distribution.

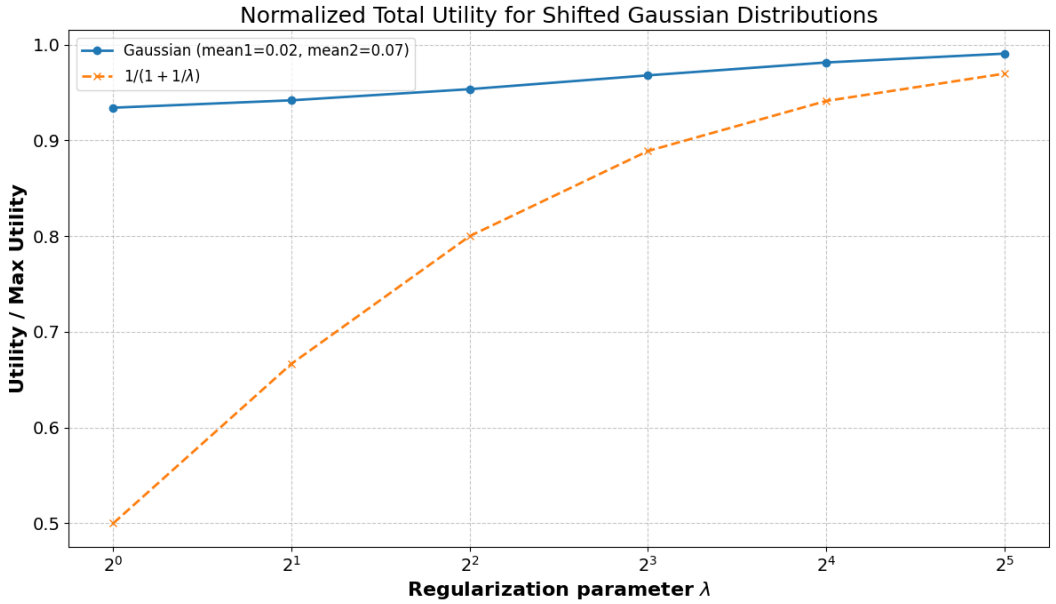


Figure 8: The price of fairness for proportional fairness with different values of λ and the theoretial lower bound, for capacity 0.01 and the shifted Gaussian distribution

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: the claims made in the abstract and introduction are followed by formal claims made in the body of the paper, and proofs in the supplementary material.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We include a limitations discussion at the end of the introduction. Throughout the paper, we explicitly state our modeling assumptions and provide justification for them. For example, we explain our choice of modeling utility via true positive count and relate it to standard loss functions. Our theorem statements also make assumptions explicit and include error terms that capture calibration error and deviations from satisfying the fairness constraints.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The theoretical results explicitly state their assumptions. The appendix also includes a section with definitions and basic claims on quantile functions, for completeness.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper is theoretical and does not include experimental results. It includes simple simulations on synthetic data to illustrate the theoretical findings. Details of these simulations are provided in the appendix and are sufficient for reproduction.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [N/A]

   Justification: The paper is theoretical and does not rely on real-world data. The simulations use simple synthetic data generated from a Gaussian distribution, with parameters (mean and variance) specified in the appendix.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper is theoretical and does not involve experiments with learning algorithms.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper is theoretical and does not involve experiments with learning algorithms.

   Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper is theoretical and does not involve experiments with learning algorithms.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have read the code of ethics and the research conforms to it.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

Justification: This work addresses fairness with the goal of promoting positive societal impact. We also discuss potential misuse of our results in the limitations section.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: Our paper is theoretical and does not pose such risks.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [N/A]

Justification: The paper is theoretical and does not use existing assets.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.

- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involves crowdsourcing.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involves crowdsourcing or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The usage of LLM done in the paper is standard, for writing, editing and searching.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.