Multiplicative Metric Fairness Under Composition

Milan Mossé ✉
Department of Philosophy, University of California at Berkeley, CA, USA

Abstract

Dwork, Hardt, Pitassi, Reingold, & Zemel [6] introduced two notions of fairness, each of which is meant to formalize the notion of similar treatment for similarly qualified individuals. The first of these notions, which we call additive metric fairness, has received much attention in subsequent work studying the fairness of a system composed of classifiers which are fair when considered in isolation [3, 4, 7, 8, 12] and in work studying the relationship between fair treatment of individuals and fair treatment of groups [6, 7, 13]. Here, we extend these lines of research to the second, less-studied notion, which we call multiplicative metric fairness. In particular, we exactly characterize the fairness of conjunctions and disjunctions of multiplicative metric fair classifiers, and the extent to which a classifier which satisfies multiplicative metric fairness also treats groups fairly. This characterization reveals that whereas additive metric fairness becomes easier to satisfy when probabilities of acceptance are small, leading to unfairness under functional and group compositions, multiplicative metric fairness is better-behaved, due to its scale-invariance.

2012 ACM Subject Classification Mathematics of computing → Probability and statistics

Keywords and phrases algorithmic fairness, metric fairness, fairness under composition

Digital Object Identifier 10.4230/LIPIcs.FORC.2023.4

Acknowledgements Many thanks to Omer Reingold and Li-Yang Tan for their generous guidance and support with this project. Thanks to James Evershed, Wes Holliday, Niko Kolodny, Gabrielle Candès, and three anonymous reviewers for helpful comments.

1 Introduction

We study the fairness of a decision-maker, modeled as a *classifier* C, which takes as input an individual and outputs a label 1 or 0, each with some probability. For example, C could take as input an individual applying for a loan and output 1 if it decides that they will receive the loan and 0 if not, and C could have high likelihood of approving application of individuals with high credit scores and a low likelihood of approving applications of individuals with low credit scores.

One plausible constraint on a fair decision-maker requires that it treat similarly qualified individuals similarly. Dwork, Hardt, Pitassi, Reingold, & Zemel [6] introduced two notions of fairness, each meant to formalize this constraint. The first of these, additive metric fairness, has received much attention in subsequent work [3, 4, 7, 8, 12, 13]:

▶ **Definition 1** (Additive metric fairness). *Let $\mathcal{U}$ denote a set of individuals. A classifier C is* additive metric fair *with respect to a metric $d : \mathcal{U} \times \mathcal{U} \to [0, 1]$ if for all $u, v \in \mathcal{U}$,*

$$|\Pr[C(u) = 1] - \Pr[C(v) = 1]| \leq d(u, v).$$

The difference in two individuals' treatment is modeled as the additive difference in their likelihoods of acceptance by the classifier, and the difference in their qualifications is given by a metric. Additive metric fairness thus requires that two individuals' difference in treatment not exceed their difference in qualifications. For example, where $\Pr[C(u) = 1]$ is the likelihood that the loan application of u is approved, $d(u, v)$ could be the normalized difference between the credit scores of u and v.

Additive metric fairness becomes easy to satisfy when the probabilities $\Pr[C(u) = 1]$ are small:

4th Symposium on Foundations of Responsible Computing (FORC 2023).
Editor: Kunal Talwar; Article No. 4; pp. 4:1–4:11
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ **Example 2** (An unfair lottery). Suppose that every pair of individuals u, v differs in qualifications by at least some amount δ. Then provided that for all individuals u, the likelihood $\Pr[C(u) = 1]$ is at most some sufficiently small value ϵ, the classifier C will be additive metric fair:

$$|\Pr[C(u) = 1] - \Pr[C(v) = 1]| \leq \max(\Pr[C(u) = 1], \Pr[C(v) = 1]) \\ \leq \epsilon \leq \delta \leq d(u, v).$$

For example, C could be a highly selective university, so that $C(u) = 1$ means that u is accepted; an investment with a low likelihood of return, so that $C(u) = 1$ means that u received a return on the investment; or a lottery, so that $C(u) = 1$ means that u had a winning lottery ticket.

As a result, additive metric fairness is compatible with the following kinds of unfairness:

▶ **Example 3** (Unfairness for groups). Suppose that there are two groups A and B of investors. If those in group B invested a cent more than those in group A, we may set $d(u, v) = .01$ for $u \in A$ and $v \in B$. A classifier C can satisfy additive metric fairness by giving those in group A no chance of receiving a return on their sizable investment while giving those in group B some sufficiently small chance ϵ of receiving a return on their similarly-sized investment. However, this is manifestly unfair to those in group A.

▶ **Example 4** (Unfairness under functional composition). Suppose u and v each apply to several universities $C_1, ..., C_k$, such that at each university C_i, the likelihood that u is accepted is 0 while $\Pr[C_i(v) = 1] = \epsilon$. Then the likelihood that v is accepted by at least one university may approach 1, while u has no chance of acceptance at any university. Even if the universities satisfy additive metric fairness when considered in isolation, because the likelihoods of acceptance are sufficiently small, they compose to create system which fails to treat similarly qualified applicants similarly.

Thus additive metric fairness is easier to satisfy when probabilities of acceptance are small, and this can lead to unfairness for groups and under functional composition. In this paper we find that the second, scale-invariant notion of fairness introduced by Dwork, Hardt, Pitassi, Reingold, & Zemel, *multiplicative metric fairness*, is better-behaved in its treatment of groups and under functional composition:

▶ **Definition 5** (Multiplicative metric fairness). *A classifier C is* multiplicative metric fair *with respect to a metric $d : \mathcal{U} \times \mathcal{U} \to \mathbb{R}^{\geq 0}$ if for all $u, v \in \mathcal{U}$,*

$$\Pr[C(u) = 1] \leq \Pr[C(v) = 1] \cdot \exp(d(u, v)).$$

Multiplicative metric fairness models the difference in treatment between two individuals not as an additive difference, but as a ratio; it does not become easy to satisfy when probabilities are small. In order to state our results, we now introduce the relevant notions of group fairness and of fairness under functional composition.

Group fairness

We propose the following notion of group fairness:

▶ **Definition 6** (Geometric Metric Fairness). *Fix a collection of protected attributes $\mathcal{A} \subseteq 2^{\mathcal{U}}$ (e.g. races, ages, genders, etc.). A classifier C satisfies* geometric metric fairness *with respect to a metric $d : \mathcal{A} \times \mathcal{A} \to \mathbb{R}^{\geq 0}$ when for all $A, B \in \mathcal{A}$,*

$$p_\Pi(A) \leq p_\Pi(B) \cdot \exp\left(d(A, B)\right),$$

where $p_\Pi(A) = \prod_{u \in A} \Pr[C(u) = 1]^{1/|A|}$ is the geometric mean likelihood of acceptance.

In the above definition, a metric quantifies differences in qualifications between groups, just as in Definitions 1 and 5, a metric quantifies differences in qualifications between individuals. For example, suppose that every job applicant $u \in A$ can be paired with some unique applicant $v \in B$ who is equally qualified, and vice versa. Then even if individual applicants within each group differ in their qualifications, there is no difference in qualifications between the groups: $d(A, B) = 0$. In this case, geometric metric fairness amounts to the constraint that

$$p_\Pi(A) = p_\Pi(B).$$

This contrasts with a well-studied notion of group fairness:

▶ **Definition 7** (Conditional Parity). *Fix $\mathcal{Q} \subseteq 2^{\mathcal{U}}$ and a collection of protected attributes $\mathcal{A} \subseteq 2^{\mathcal{U}}$ (e.g. races, ages, genders, etc.). A classifier C satisfies* conditional parity *if for all $Q \in \mathcal{Q}, A, B \in \mathcal{A}$,*

$$p_\Sigma(A \cap Q) = p_\Sigma(B \cap Q)$$

where $p_\Sigma = \frac{1}{|A \cap Q|} \sum_{u \in A \cap Q} \Pr[C(u) = 1]$ is the arithmetic mean likelihood of acceptance.

Conditional parity was introduced by Ritov, Sun, & Zhao [14] and plays a central role in Dwork & Ilvento's study of fairness under composition [7]. Conditional parity generalizes other group notions of fairness. For example, one recovers *parity* by setting $\mathcal{Q} = \{\mathcal{U}\}$; one recovers *equalized odds* by setting $\mathcal{Q} = \{\{u : Y(u) = y\} : y \in \{0, 1\}\}$, where $Y(u)$ denotes the true label of u; and one recovers *equal opportunity* by setting $\mathcal{Q} = \{\{u : Y(u) = 1\}\}$ [11, 15]. In general, we think of $\mathcal{Q}$ as a collection of sets of individuals who are similarly qualified for the purposes of classification.

Plausibly, one should not be able to "make up for" mistreatment of some individuals within a group by treating other individuals within the group better; a radical departure from the mean treatment for any sub-group should register as unfair. However, because conditional parity only constrains the arithmetic mean probability of acceptance across members of a group, it allows for large variance in treatment of individuals within a group. In 2010, this feature of the arithmetic mean led the United Nations to change its way of calculating the Human Development Index (HDI):

> In 2010, the geometric mean was introduced to compute the HDI [which was previously computed with the arithemtic mean]. Poor performance in any dimension is directly reflected in the geometric mean. In other words, a low achievement in one dimension is not linearly compensated for by a higher achievement in another dimension. The geometric mean reduces the level of substitutability between dimensions and at the same time ensures that a 1 percent decline in the index of, say, life expectancy has the same impact on the HDI as a 1 percent decline in the education or income index. Thus, as a basis for comparisons of achievements, this method is also more respectful of the intrinsic differences across the dimensions than a simple average.

Just as the geometric mean index value is thought to better respect differences and non-substitutability across the dimensions of the HDI, the geometric mean likelihood of acceptance across a group might be thought to also better respect differences and non-substituability across individuals within a group; this motivates the constraint of geometric metric fairness.[1]

[1] When any factor of the geometric mean is 0, of course the geometric mean is itself 0, and it becomes trivial to ensure geometric metric fairness; one merely has to assign $\Pr[C(u) = 1] = \Pr[C(v) = 1] = 0$ for one person $u \in A$ and another person $v \in B$. For this reason, the geometric mean (and the associated definition of fairness) is most meaningful when probabilities are nonzero.

Fairness under functional composition

We focus on the two kinds of functional composition introduced (with the following examples) by Dwork & Ilvento [7]:

- *AND.* Suppose that, considered in isolation from one another, a university's admissions and financial aid committees treat every pair of similarly qualified prospective students similarly. To what degree do similarly qualified students have similar likelihoods of receiving admission *and* financial aid offers?
- *OR.* Suppose that, considered in isolation from one another, several universities' admissions committees treat every pair of similarly qualified prospective students similarly. To what degree do two similarly qualified individuals have very different overall likelihoods of being accepted by at least one university?

More formally, we can define the AND and OR compositions of several classifiers:

▶ **Definition 8.** *Fix classifiers $C_1, ..., C_k$. Where u is an individual to be classified, define the classifiers*

$$C_{AND}(u) = \bigwedge_{i \in [k]} C_i(u)$$
$$C_{OR}(u) = \bigvee_{i \in [k]} C_i(u).$$

In other words, C_{AND} accepts individual u if and only if each of $C_1, ..., C_k$ accepts u, and C_{OR} accepts u if and only if at least one of the classifiers $C_1, ..., C_k$ accepts u.
Supposing C_i is (additive or multiplicative) metric fair with respect to $d_i(u, v)$ for $i \in [k]$, in fairness under functional composition, we ask: With respect to what metric are C_{AND} and C_{OR} (additive or multiplicative) metric fair?

Our results

Having introduced the relevant definitions, we can state the paper's results:

▶ **Theorem 9** (Groups)**.** *If C is multiplicative metric fair with respect to d, then it is geometric metric fair with respect to $\mathsf{EMD}_d(A, B)$, the earth-mover distance between uniform distributions on A and B, with $d(u, v)$ giving the cost of moving a unit of probability from u to v. Further, this is tight: for any metric d on $\mathcal{U}$, there exists a classifier C which is multiplicative metric fair with respect to d and for which*

$$p_\Pi(A) = p_\Pi(B) \cdot \exp\left(\mathsf{EMD}_d(A, B)\right).$$

▶ **Theorem 10** (Functions)**.** *If C_i is multiplicative metric fair with respect to $d_i(u, v)$ for $i \in [k]$, then:*

- *C_{AND} is multiplicative metric fair with respect to $d_\Sigma(u, v) = \sum_i d_i(u, v)$.*
- *C_{OR} is multiplicative metric fair with respect to $d_{\max}(u, v) = \max_i d_i(u, v)$.*

Further, these results are tight: for each of the above forms of composition and for any choices of d_i for $i \in [k]$, there exist classifiers C_i for $i \in [k]$ which are multiplicative metric fair with respect to d_i and whose composition is multiplicative metric fair with respect to no metric smaller than the one stated above.

The rest of the paper is organized as follows. §1.1 discusses related work. §1.2 summarizes the notation used in the paper. §2 discusses the relationship between notions of metric fairness and the above notions of group fairness, proving the paper's first main result. §3 discusses how notions of metric fairness behave under functional composition, proving the paper's second main result.

1.1 Related work

Several recent works in algorithmic fairness studies how the fairness of classifiers locally relates to that of the classifiers' global behavior composed over many decisions, or to that of a classifier that in some way composes the decisions of the individual classifiers [6, 2, 7, 9, 10, 8, 12, 3, 4].

The need for work in this area is underscored by the fact that in practice, classifiers are often trained separately and without communication, so that any guarantees on their global behavior must rest solely on the decisions the designers of the classifiers are able to make in isolation. In a recent survey of work on fairness in machine learning, Chouldechova and Roth make exactly this point, calling for work exploring fairness under composition [5]:[2]

> Experience from differential privacy suggests that graceful degradation under composition is key to designing complicated algorithms satisfying desirable statistical properties, because it allows algorithm design and analysis to be modular. Thus, it seems important to find satisfying fairness definitions and richer frameworks that behave well under composition.

Much recent work on fairness under composition focuses in particular on the behavior of additive metric fairness under various kinds of composition [7, 8, 12, 3]. There are two papers which relate especially closely to this one. The first, written by Dwork, Hardt, Pitassi, Reingold, & Zemel [6], introduced the notion of additive metric fairness and characterized its relation to conditional parity. The second, written by Dwork & Ilvento [7], introduced the kinds of functional composition (AND and OR) studied in this paper and made progress in showing that additive metric fairness is not always well-behaved under these kinds of composition; we summarize some of this work in §3.1. Our results are meant to complement this line of research, by showing that multiplicative metric fairness is better-behaved in treatment of groups and under functional composition.

1.2 Notation

A classifier $C : \mathcal{U} \times \{0,1\}^* \to \{0,1\}$ is a (possibly randomized) Boolean-valued map, defined on a universe $\mathcal{U}$ of individuals $u \in \mathcal{U}$. We denote by $1 - C$ the classifier C' which accepts precisely the individuals rejected by C: $C'(u) := \neg C(u)$. We say that a classifier *accepts* an individual when it assigns them a label of 1 and *rejects* an individual when it assigns them a label of 0. Throughout, d is a metric on $\mathcal{U}$. For classifiers $C_1, ..., C_k$, we use $d_1, ..., d_k$ to denote their corresponding metrics. We define $d_\Sigma(u, v) = \sum_i d_i(u, v)$ and $d_{\max}(u, v) = \max_i d_i(u, v)$.

Subsets $A, B \subseteq \mathcal{U}$ denote protected groups. For a classifier C, we denote by $p_\Pi(A)$ the geometric mean likelihood of acceptance and by $p_\Sigma(A)$ the arithmetic mean likelihood of acceptance. When there are multiple classifiers $C_1, ..., C_k$, we assume the randomness of the classifiers C_i to be mutually independent, and we use $p_i(u)$ to denote $\Pr[C_i(u) = 1]$. We define $p_{AND}(u) = \Pr[C_{AND}(u) = 1]$ and $p_{OR}(u) = \Pr[C_{OR}(u) = 1]$, where C_{AND} is the AND composition and C_{OR} the OR composition of some classifiers $C_1, ..., C_k$.

[2] Dwork & Ilvento [7] point out an important difference between differential privacy and fairness under composition: "Comparing functional composition to differential privacy, it is important to understand that each component satisfying individual fairness separately (and for different metrics) is not analogous to the composition properties of differential privacy. With differential privacy, we assume a single privacy loss random variable which evolves gracefully with each release of information, increasing in expectation over time. However, with fairness, we may see that fairness loss increases or decreases (depending on the number and type of compositions) in idiosyncratic ways. Moreover, we may need to simultaneously satisfy many different task-specific 'fairness budgets,' and a bounded increase in distance based on one task may be catastrophically large for another."

2 Treatment of groups

The relation of (conditional) parity to additive metric fairness has garnered recent interest [1, 6].[3] Dwork, Hardt, Pitassi, Reingold, & Zemel [6] give a tight characterization of the relationship between additive metric fairness and parity, using the following notion of earth-mover's distance:

▶ **Definition 11** (Earth-Mover's Distance). *Fix sets $A, B \subseteq \mathcal{U}$ and a collection of associated costs $d(u, v) \geq 0$ for each $u \in A, v \in B$. The earth-mover's distance $\mathsf{EMD}_d(A, B)$ is the minimum amount of work required to transform a uniform distribution on A into one on B, where the amount of work required to move a unit of probability from individual u to individual v is given by $d(u, v)$. Formally,*

$$\mathsf{EMD}_d(A, B) = \sum_{u \in A, v \in B} f_{u,v} \cdot d(u, v),$$

where the variables $f_{u,v}$ give an optimal solution to the following linear program (LP):

$$\min \sum_{u \in A, v \in B} f_{u,v} \cdot d(u, v)$$

$$f_{u,v} \geq 0$$

$$f_{u,v} = 0 \text{ if } u \notin A \text{ or } v \notin B$$

$$\sum_{v \in B} f_{u,v} = \frac{1}{|A|}, \sum_{u \in A} f_{u,v} = \frac{1}{|B|}, \sum_{u \in A, v \in B} f_{u,v} = 1$$

Dwork, Hardt, Pitassi, Reingold, & Zemel [6] prove the following by LP duality, applied to the LP in Definition 11:

▶ **Theorem 12.** *If C is additive metric fair with respect to d, then for all $A, B \subseteq \mathcal{U}$,*

$$|p_\Sigma(A) - p_\Sigma(B)| \leq \mathsf{EMD}_d(A, B).$$

Further, this is tight: for all metrics d and choices of A, B, there exists a classifier C that is additive metric fair with respect to d, such that the above inequality is an equality.

The above result says that if C is additive metric fair, then the earth-mover distance gives a tight characterization of the extent to which C satisfies conditional parity. The same authors observe that an identical upper bound holds if we instead assume that C and $1 - C$ are multiplicative metric fair:

▶ **Corollary 13.** *If C and $1 - C$ are multiplicative metric fair with respect to d, then for all $A, B \subseteq \mathcal{U}$,*

$$|p_\Sigma(A) - p_\Sigma(B)| \leq \mathsf{EMD}_d(A, B).$$

[3] We observe that there is no deep difference been parity (i.e. conditional parity where $\mathcal{Q} = \mathcal{U}$) and conditional parity. It is clear that conditional parity is a generalization of parity. Conversely, conditional parity is a version of parity where we stipulate that A and B are equally qualified.

Thus the relationship between additive metric fairness and parity is well-understood, and it is known already that multiplicative metric fairness performs "at least as well" as additive metric fairness, in the sense that one can only get closer to satisfying parity in the multiplicative case.

However, because Theorem 12 only provides a bound on the difference between *arithmetic mean* conditional probabilities of acceptance, the guarantee can still hold when sub-groups are treated very differently, so long as advantages and disadvantages of different sub-groups are traded off in a way that maintains conditional parity. We now show Theorem 9, according to which multiplicative metric fairness, in contrast to additive metric fairness, provides a bound on the ratio of the geometric mean probabilities of acceptance:

Proof. We first show the upper bound and next show the lower bound. Fix a classifier C which is multiplicative metric fair with respect to d, and fix any flow $\{f_{u,v}\}_{u\in A, v\in B}$ solving the earth-mover LP. Using the multiplicative metric fairness constraint, note that for all $u \in A, v \in B$, we have

$$\Pr[C(u) = 1]^{f_{u,v}} \leq e^{d(u,v)\cdot f_{u,v}} \cdot \Pr[C(v) = 1]^{f_{u,v}}.$$

Taking the product on both sides over all $u \in A, v \in B$ gives

$$\prod_{u\in A} \Pr[C(u) = 1]^{\sum_{v\in B} f_{u,v}} \leq \prod_{u\in A, v\in B} e^{d(u,v)\cdot f_{u,v}} \cdot \prod_{v\in B} \Pr[C(v) = 1]^{\sum_{u\in A} f_{u,v}}.$$

Note that for $u \in A$, we have $\sum_{v\in B} f_{u,v} = 1/|A|$, while for $v \in B$, we have $\sum_{u\in A} f_{u,v} = 1/|B|$. Thus

$$p_\Pi(A) \leq \exp\{\mathsf{EMD}_d(A, B)\} \cdot p_\Pi(B)$$

Now, we show the lower bound. Fix any metric d. Let c be a constant with $1 \leq c$ such that $d(u, v) \leq c$ for all $u, v \in \mathcal{U}$. Define the metric $d'(u, v) = d(u, v)/c$, so that $d'(u, v) \in [0, 1]$ for all $u, v \in \mathcal{U}$. By Theorem 12, there exists a classifier C' which is additive metric fair with respect to d' and for which

$$|p'_\Sigma(A) - p'_\Sigma(B)| = \mathsf{EMD}_{d'}(A, B). \tag{1}$$

Define C by

$$\Pr[C(u) = 1] = \exp\{-\Pr[C'(u) = 1] \cdot c\}.$$

Because C' is additive metric fair with respect to d', it follows that C is multiplicative metric fair with respect to d:

$$\ln\left[\frac{\Pr[C(u) = 1]}{\Pr[C(v) = 1]}\right] = -\Pr[C'(u) = 1] \cdot c + \Pr[C'(v) = 1] \cdot c \leq d'(u, v) \cdot c = d(u, v).$$

Suppose without loss of generality that $p'_\Sigma(B) \geq p'_\Sigma(A)$. Let us restate Equation 1 using the definition of C:

$$\begin{aligned}\frac{1}{c} \cdot \sum_{u\in B} \frac{-\ln \Pr[C(u) = 1]}{|B|} - \frac{1}{c} \cdot \sum_{v\in A} \frac{-\ln \Pr[C(v) = 1]}{|A|} &= \mathsf{EMD}_{d'}(A, B) \\ &= \frac{\mathsf{EMD}_d(A, B)}{c}.\end{aligned}$$

Eliminating the factor $1/c$ and making both sides the exponent of e, we obtain

$$\frac{\prod_{u\in B} \Pr[C(u) = 1]^{-1/|B|}}{\prod_{u\in A} \Pr[C(u) = 1]^{-1/|A|}} = \frac{p_\Pi(A)}{p_\Pi(B)} = \exp\{\mathsf{EMD}_d(A, B)\},$$

as desired. ◀

3 Functional composition

We first overview known results for additive metric fairness under functional composition; this will serve to illustrate the contrast with multiplicative metric fairness.

3.1 Additive metric fairness under functional composition

Here, we rehearse known limitations and positive results for additive metric fairness of AND and OR compositions, with an eye to explaining some of the difficulties that arise.

We start with AND fairness. Given the following result, it is tempting to conjecture that C_{AND} is additive metric fair with respect to the maximum of the individual metrics:

▶ **Proposition 14** (Dwork & Ilvento [7]). *Fix nontrivial metrics d_1, d_2 and let d be any metric. If there exist $u, v \in \mathcal{U}$ such that*

- $d(u,v) \leq d_1(u,v), d_2(u,v)$*, and*
- $d_1(u,v), d_2(u,v) > 0$*,*

there exist C_1, C_2, fair with respect to d_1, d_2, such that C_{AND} is unfair with respect to d.

dxBut in fact even picking $d_{\max}$ does not guarantee additive metric fairness:

▶ **Example 15.** Let C_1 and C_2 be copies of the same classifier: $p_i(u) = 1$, $p_i(v) = 1/2$, and $d_i(u,v) = 1/2$ for $i = 1, 2$. Then the classifiers C_i are individually additive metric fair with respect to $d_i(u,v)$, but their composition is not fair with respect to $d_{\max}(u,v) = \max_i d_i(u,v)$.

In a sense, when probabilities are small, the choice of metric for the AND composition in the additive case does not matter: as long as for each u, there exists some i with $p_i(u) \leq d(u,v)$, a fortiori $p_{AND}(u) \leq d(u,v)$. Since without loss of generality $p_{AND}(u) \geq p_{AND}(v)$, we have $|p_{AND}(u) - p_{AND}(v)| \leq d(u,v)$, giving fairness with respect to the arbitrary metric d. In other words, if probabilities are small enough, additive metric fairness for the AND composition trivializes.

We turn now to OR fairness. Dwork & Ilvento [7] observe that in the case of OR fairness, it is natural to suppose that the metrics are identical; returning to an earlier example, if the individual classifiers are admissions committees for different universities, it is natural to suppose that the admissions committees compare candidates using similar metrics. In this case the problem just discussed of picking a metric against which to compare the composition is more tractable: one can pick the composition metric to be the same as the metrics of the individual classifiers. Dwork & Ilvento's results imply the following:

▶ **Proposition 16** (Dwork & Ilvento [7]). *Fix classifiers $C_1, ..., C_k$ that are additive metric fair with respect to d. Consider two cases. If for all u, we have*

$$\Pr[C_{OR}(u) = 1] \geq \frac{1}{2},$$

then for any classifier C_{k+1} with $\Pr[C_{k+1}(u) \geq 1/2]$ for all $u \in \mathcal{U}$, the OR composition of $C_1, ..., C_{k+1}$ is additive metric fair with respect to d. If instead the above condition fails for some u, v with nontrivial distance ($d(u,v) > 0$), then there exist two classifiers C_{k+1}, C_{k+2}, additive metric fair with respect to d, such that the OR composition of $C_1, ..., C_{k+2}$ is not additive metric fair with respect to d.

In other words, the first, positive part of the above result says that if an initial collection of classifiers is more likely than not to accept every individual, adding a classifier that shares this property makes the entire collection's OR composition fair. The second, negative part of the result says that if there are even two (nontrivially different) individuals the initial collection is more likely to reject than accept, one can add two fair classifiers that make the OR composition of the entire collection unfair. We earlier found that when when the probabilities $p_i(u)$ are small enough, additive fairness for the AND composition trivializes; we now find that when the probabilities are small, we have no positive result for the additive metric fairness of the OR composition.

3.2 Multiplicative metric fairness under functional composition

We now show Theorem 10, which provides substantive fairness guarantees even when probabilities of acceptance are small:[4]

Proof. For $u, v \in \mathcal{U}$

$$p_{AND}(u) = \prod_i p_i(u) \leq \prod_i p_i(v) \cdot e^{d_i(u,v)} = p_{AND}(v) \cdot e^{\sum_i d_i(u,v)}.$$

This shows that C_{AND} is multiplicative metric fair with respect to d_Σ. To see that the result is tight, one simply picks classifiers such that $p_i(u) = e^{d_i(u,v)} \cdot p_i(v)$, so that indeed $p_{AND}(u) = e^{d_\Sigma(u,v)} p_{AND}(v)$.

We next show that C_{OR} is multiplicative metric fair with respect to $d_{\max}$ and show that this is tight. We only consider the case where $k = 2$, since iterating the argument then gives the result for general k. Suppose without loss of generality that $p_{OR}(u) \geq p_{OR}(v)$. By assumption C_1, C_2 are multiplicative metric fair with respect to $d_{\max}$, so it suffices to show the first inequality:

$$\frac{p_{OR}(u)}{p_{OR}(v)} \leq \max\left[\frac{p_1(u)}{p_1(v)}, \frac{p_2(u)}{p_2(v)}\right] \leq e^{d_{\max}(u,v)}.$$

To show the first inequality, we suppose $p_{OR}(u)/p_{OR}(v) > p_1(u)/p_1(v)$ and show that it follows that

$$\frac{p_{OR}(u)}{p_{OR}(v)} < \frac{p_2(u)}{p_2(v)}.$$

Noting that $p_{OR}(u) = p_1(u) + p_2(u) - p_1(u)p_2(u)$, let us rephrase $\frac{p_{OR}(u)}{p_{OR}(v)} > p_1(u)/p_1(v)$ after clearing denominators:

$$p_1(v)[p_1(u) + p_2(u) - p_1(u)p_2(u)] > p_1(u)[p_1(v) + p_2(v) - p_1(v)p_2(v)].$$

After removing $p_1(v)p_1(u)$ from both sides and factoring, the above says that

$$p_2(u)p_1(v)(1 - p_1(u)) > p_2(v)p_1(u)(1 - p_1(v)).$$

In other words,

$$\frac{p_2(u)}{p_2(v)} > \frac{p_1(u)}{p_1(v)} \cdot \frac{1 - p_1(v)}{1 - p_1(u)}.$$

[4] Dwork, Hardt, Pitassi, Reingold & Zemel [6] introduce the constraint equivalent to multiplicative metric fairness for C and $1 - C$. Theorem 10 illustrates why this paper has separated their definition into two components: multiplicative metric fairness of $1 - C_i$ for $i \in [k]$ does not yield multiplicative metric fairness for $1 - C_{AND}$, where C_{AND} is the AND composition of $C_1, ..., C_k$, but instead yields multiplicative metric fairness for the AND composition of $1 - C_1, ..., 1 - C_k$.

We will later show that $\frac{1-p_1(v)}{1-p_1(u)} \geq \frac{1-p_2(u)}{1-p_2(v)}$, but let us finish the proof on this assumption. Combining this with the above inequality gives

$$\frac{p_2(u)}{p_2(v)} > \frac{p_1(u)}{p_1(v)} \cdot \frac{1-p_2(u)}{1-p_2(v)}.$$

Clearing denominators, the above says that

$$p_2(u)[p_1(v) - p_1(v)p_2(v)] > p_2(v)[p_1(u) - p_1(u)p_2(u)]$$

Add $p_2(u)p_2(v)$ to both sides. Then the above says that

$$p_2(u)[p_1(v) + p_2(v) - p_1(v)p_2(v)] > p_2(v)[p_1(u) + p_2(u) - p_1(u)p_2(u)],$$

or in other words, $p_2(u)/p_2(v) > p_{OR}(u)/p_{OR}(v)$, as desired.

It remains for us to show that

$$\frac{1-p_1(v)}{1-p_1(u)} \geq \frac{1-p_2(u)}{1-p_2(v)}.$$

Since by assumption $p_{OR}(u) \geq p_{OR}(v)$, of course $1 - p_{OR}(v) \geq 1 - p_{OR}(u)$. Noting that $p_{OR}(v) = 1 - (1 - p_1(v))(1 - p_2(v))$, we can rephrase $1 - p_{OR}(v) \geq 1 - p_{OR}(u)$ as

$$(1 - p_1(v))(1 - p_2(v)) \geq (1 - p_1(u))(1 - p_2(u)) \iff \frac{1-p_1(v)}{1-p_1(u)} \geq \frac{1-p_2(u)}{1-p_2(v)}.$$

We now show that the result for OR is tight. It again suffices to consider the case for $k = 2$. Fix any metric $d_1(u, v)$ and put $d_2(u, v) = d_1(u, v) - \alpha$ for an arbitrarily small $\alpha > 0$. We claim there exist classifiers C_1, C_2 such that:

- The classifiers C_1, C_2 are (respectively) multiplicative metric fair with respect to d_1, d_2.
- C_{OR} is not multiplicative metric fair with respect to $d_2(u, v)$.

Let $\beta_1 \in (0, e^{-d_1(u,v)}], \beta_2 \in (0, e^{-d_2(u,v)}]$ be parameters to be chosen later and define

$$p_1(u) = \beta_1 \cdot \exp[d_1(u, v)]$$
$$p_1(v) = \beta_1$$
$$p_2(u) = \beta_2 \cdot \exp[d_2(u, v)]$$
$$p_2(v) = \beta_2.$$

Then the classifiers C_1, C_2 defined by the above probabilities are multiplicative metric fair with respect to d_1, d_2 (respectively). We claim that for $\beta_2 < \frac{\beta_1 \cdot \alpha}{\exp[d_2(u,v)] \cdot (1-\beta_1)}$, we have

$$p_{OR}(u) \geq p_1(u) > \exp[d_2(u, v)] \cdot p_{OR}(v),$$

so that C_{OR} is indeed not multiplicative metric fair with respect to d_2. It suffices to show the inequality on the right, which says that

$$p_1(u) > \exp[d_2(u, v)][\beta_1 + \beta_2 - \beta_1\beta_2] = \exp[d_2(u, v)] \cdot \beta_1 + \beta_2 \cdot \exp[d_2(u, v)] \cdot (1 - \beta_1).$$

Subtracting $\exp[d_2(u, v)] \cdot \beta_1$ from both sides, this says that

$$\beta_1 \cdot \alpha = p_1(u) - \exp[d_2(u, v)] \cdot \beta_1 > \beta_2 \cdot \exp[d_2(u, v)] \cdot (1 - \beta_1),$$

which holds by our choice of β_2. ◀

References

1 Reuben Binns. On the apparent conflict between individual and group fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 514–524, 2020.

2 Amanda Bower, Sarah N Kitchen, Laura Niss, Martin J Strauss, Alexander Vargas, and Suresh Venkatasubramanian. Fair pipelines. *arXiv preprint*, 2017. `arXiv:1707.00391`.

3 Shuchi Chawla and Meena Jagadeesan. Fairness in ad auctions through inverse proportionality. *arXiv preprint*, 2020. `arXiv:2003.13966`.

4 Shuchi Chawla, Rojin Rezvan, and Nathaniel Sauerberg. Individually-fair auctions for multi-slot sponsored search. In *3rd Symposium on Foundations of Responsible Computing*, page 1, 2022.

5 Alexandra Chouldechova and Aaron Roth. The frontiers of fairness in machine learning. *arXiv preprint*, 2018. `arXiv:1810.08810`.

6 Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226, 2012.

7 Cynthia Dwork and Christina Ilvento. Fairness under composition. *arXiv preprint*, 2018. `arXiv:1806.06122`.

8 Cynthia Dwork, Christina Ilvento, and Meena Jagadeesan. Individual fairness in pipelines. *arXiv preprint*, 2020. `arXiv:2004.05167`.

9 Vitalii Emelianov, George Arvanitakis, Nicolas Gast, Krishna Gummadi, and Patrick Loiseau. The price of local fairness in multistage selection. *arXiv preprint*, 2019. `arXiv:1906.06613`.

10 Swati Gupta and Vijay Kamble. Individual fairness in hindsight. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 805–806, 2019.

11 Moritz Hardt, Eric Price, and Nati Srebro. Equality of opportunity in supervised learning. In *Advances in neural information processing systems*, pages 3315–3323, 2016.

12 Christina Ilvento, Meena Jagadeesan, and Shuchi Chawla. Multi-category fairness in sponsored search auctions. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 348–358, 2020.

13 Michael P Kim, Omer Reingold, and Guy N Rothblum. Fairness through computationally-bounded awareness. *arXiv preprint*, 2018. `arXiv:1803.03239`.

14 Ya'acov Ritov, Yuekai Sun, and Ruofei Zhao. On conditional parity as a notion of non-discrimination in machine learning. *arXiv preprint*, 2017. `arXiv:1706.08519`.

15 Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rogriguez, and Krishna P Gummadi. Fairness constraints: Mechanisms for fair classification. In *Artificial Intelligence and Statistics*, pages 962–970. PMLR, 2017.