A SUPPLEMENTARY MATERIAL

A.1 NUMERICAL STABILITY

A.1.1 VON MISES-FISHER CONSTANTS

Recall the loss defined in Equation 7:

$$\mathcal{L}_{\text{FvMF}} = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{C_d(\kappa_{a_{y_i}})e^{\kappa_{a_{y_i}}\boldsymbol{\mu}_{y_i}^\intercal \boldsymbol{z}_i}}{\sum_{k=1}^{K} C_d(\kappa_{a_k})e^{\kappa_{a_k}\boldsymbol{\mu}_k^\intercal \boldsymbol{z}_i}}\right) \quad \text{with} \quad C_d(\kappa) = \frac{\kappa^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(\kappa)}.$$

I_ν stands for the modified Bessel function of the first kind at order ν, whose logarithm can be computed with high precision using a Python library for arbitrary-precision floating-point arithmetic such as `mpmath` (Johansson et al. (2021); Kim (2021)).

Once $\log(I_{\frac{d}{2}-1}(\kappa))$ is obtained, one is able to compute the logarithm of $C_d(\kappa)$ as:

$$\log(C_d(\kappa)) = (\frac{d}{2}-1)\log(\kappa) - \frac{d}{2}\log(2\pi) - \log(I_{\frac{d}{2}-1}(\kappa)).$$

Figure 5 displays $\log(I_{\frac{d}{2}-1}(\kappa))$ and $\log(C_d(\kappa))$ as functions of κ for $d = 512$.

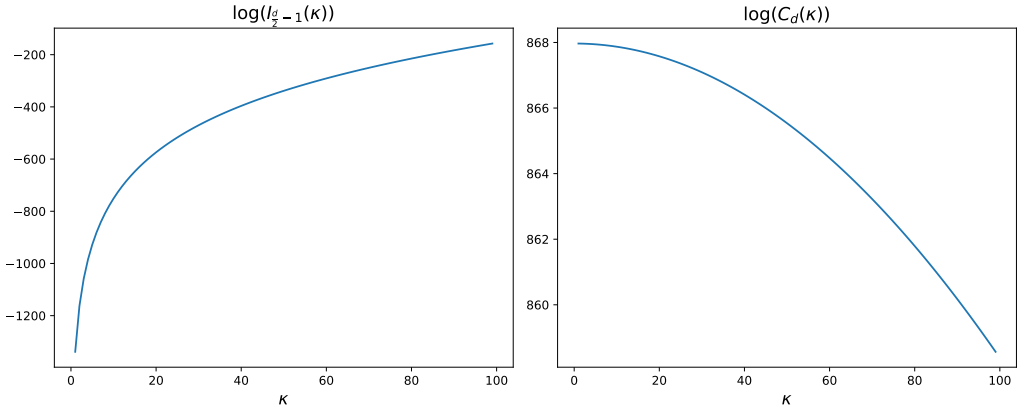


Figure 5: $\log(I_{\frac{d}{2}-1}(\kappa))$ and $\log(C_d(\kappa))$ as functions of κ for $d = 512$.

A.1.2 LOSS STABILITY

To make use of the numerical stability of the quantity $\log(C_d(\kappa))$, $\mathcal{L}_{\text{FvMF}}$ can be written as:

$$\mathcal{L}_{\text{FvMF}} = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{e^{\log(C_d(\kappa_{a_{y_i}}))+\kappa_{a_{y_i}}\boldsymbol{\mu}_{y_i}^\intercal \boldsymbol{z}_i}}{\sum_{k=1}^{K} e^{\log(C_d(\kappa_{a_k}))+\kappa_{a_k}\boldsymbol{\mu}_k^\intercal \boldsymbol{z}_i}}\right).$$

Recall the cross-entropy loss $\mathcal{L}_{CE}(\{q_{i,k}\},\{y_i\})$ with $1 \le i \le n$ and $1 \le k \le K$ defined as:

$$\mathcal{L}_{CE}(\{q_{i,k}\},\{y_i\}) = -\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{e^{q_{i,y_i}}}{\sum_{k=1}^{K} e^{q_{i,k}}}\right)$$

$\mathcal{L}_{\text{FvMF}}$ can be expressed as the cross-entropy loss:

$$\mathcal{L}_{\text{FvMF}} = \mathcal{L}_{CE}(\{q_{i,k}\},\{y_i\})$$

where the logits $q_{i,k} = \log(C_d(\kappa_{a_k})) + \kappa_{a_k}\boldsymbol{\mu}_k^\intercal \boldsymbol{z}_i$ satisfy ($\boldsymbol{\mu}_k, \boldsymbol{z}_i \in \mathbb{S}^{d-1}$):

$$\log(C_d(\kappa_{a_k})) - \kappa_{a_k} \le q_{i,k} \le \log(C_d(\kappa_{a_k})) + \kappa_{a_k}$$

Those bounds are displayed in Figure 6. The fact that $\mathcal{L}_{\text{FvMF}}$ can be expressed as the cross-entropy loss makes it possible to use the logsoftmax trick and thus further increases its numerical stability.

We provide on Figure 7 the behavior of our $\mathcal{L}_{\text{FvMF}}$ training loss, used to train the Ethical Module on top of ArcFace ResNet50.

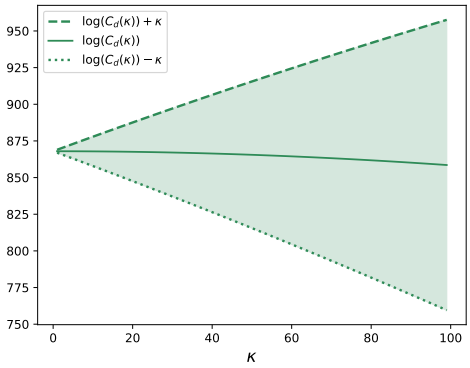


Figure 6: Range of values of the $\mathcal{L}_{\text{FvMF}}$ loss logits for $d = 512$.

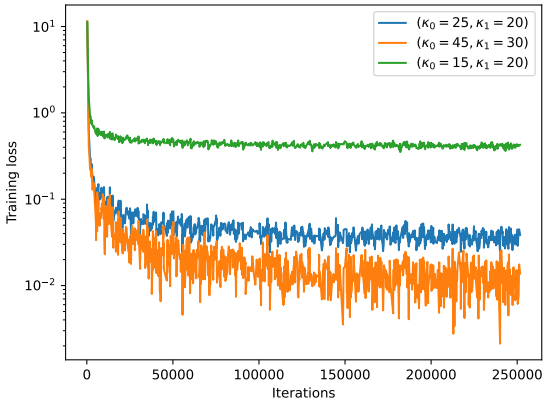
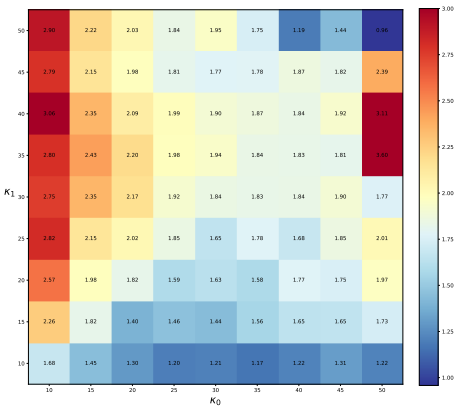


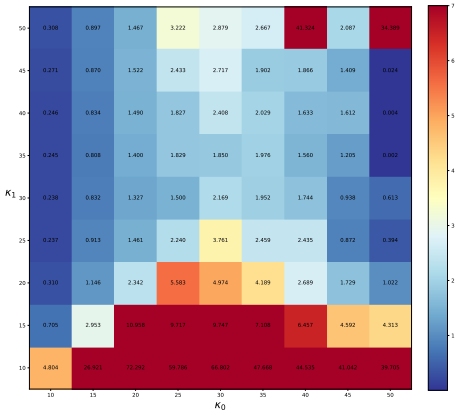
Figure 7: $\mathcal{L}_{\text{FvMF}}$ training loss for the pre-trained model ArcFace ResNet 50.

A.2 GRID-SEARCH ON IJB-C

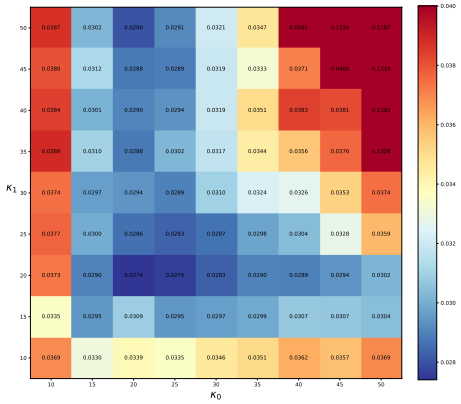
In order to select relevant pairs of gender-hyperparameters (κ_0, κ_1), we perform a grid-search on a square of size 9×9 and keep track of the canonical performance metric $\mathrm{FRR@(FAR} = 10^{-3})$ together with variants of our two fairness metrics $\mathrm{BFRR}(10^{-3})$ and $\mathrm{BFAR}(10^{-3})$ introduced in Eq. 4 and 5. These variants are respectively $\mathrm{FRR}_1(t)/\mathrm{FRR}_0(t)$ and $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$ computed at the threshold t satisfying $\max_{a\in\{0,1\}} \mathrm{FAR}_a(t) = 10^{-3}$. In this way we can visualize the inversion of bias incurred by our model: in most settings, females are disadvantaged while some extreme values of the hyperparameters disadvantage males. The results displayed in Figure 8 contain the results of Figure 4 but they are more complete.



(a) $\mathrm{FRR}_1(t)/\mathrm{FRR}_0(t)$



(b) $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$



(c) $\mathrm{FRR@(FAR} = 10^{-3})$

Figure 8: Three metrics along the grid-search. Notice that 8a and 8b are computed at the threshold t satisfying $\max_{a\in\{0,1\}} \mathrm{FAR}_a(t) = 10^{-3}$. The pre-trained model is ArcFace with a ResNet100 backbone and the Ethical Module is evaluated on IJB-C.

A.3 TWO HEURISTICS ON THE HYPERPARAMETERS TRENDS

A.3.1 FINDING AN INITIAL POINT

Assume that the identities of the training dataset are made of n_0 males and n_1 females. Our ideal goal is to find the optimal concentration parameters κ_0 and κ_1 without relying on a greed search methodology which is computationally expensive. At a heuristic level, we would like to equalize the *volume covered* by male with the *volume covered* by female in the latent hypersphere. The first

step is to properly define the volume covered by an individual. Recall the vMF loss assumes that pictures of a given identity are i.i.d. realizations of a vMF probability measure with parameters (μ, κ). It guides us to introduce the following definition of the occupancy set of an individual.

Definition (Occupancy set)**.** Let $\alpha \in [0, 1]$. The α-occupancy set $\mathcal{S}_\alpha(\mu, \kappa)$ of an individual with parameters (μ, κ) is defined by

$$\mathcal{S}_\alpha(\boldsymbol{\mu}, \kappa) := \left\{x \in \mathbb{S}^d,\, \langle \boldsymbol{x}, \boldsymbol{\mu} \rangle \leq \theta_\alpha(\kappa)\right\},$$

where $\theta_\alpha(\kappa)$ is defined by

$$\int_{\boldsymbol{x} \in \mathbb{S}^d} V_d(\boldsymbol{x}; \boldsymbol{\mu}, \kappa) \mathbf{1}_{\langle \boldsymbol{x}, \boldsymbol{\mu} \rangle \leq \theta_\alpha(\kappa)} \mathrm{d}\boldsymbol{x} = \alpha. \tag{8}$$

In words, it is the hyperspherical cap whose vMF mass is α. The area of $\mathcal{S}_\alpha(\boldsymbol{\mu}, \kappa)$ does not depend on $\boldsymbol{\mu}$ and is given by the following formula:

$$\mathcal{A}_\alpha(\kappa) = \frac{1}{2} \frac{2\pi^{d/2}}{\Gamma(d/2)} I_{\sin^2(\theta_\alpha)} \left(\frac{d-1}{2}, \frac{1}{2}\right),$$

where I is the regularized incomplete beta function.

Coming back to our initial problem, we define the α-occupancy sets of males and females by:

$$\mathcal{S}_\alpha^{(0)} := \bigcup_{1 \leq k \leq K:\ a_k = 0} \mathcal{S}_\alpha(\boldsymbol{\mu}_k, \kappa_0) \quad \text{and} \quad \mathcal{S}_\alpha^{(1)} := \bigcup_{1 \leq k \leq K:\ a_k = 1} \mathcal{S}_\alpha(\boldsymbol{\mu}_k, \kappa_1).$$

Our ideal goal is to find (κ_0^*, κ_1^*) that solves the following minimization problem:

$$\min_{\kappa_0, \kappa_1 > 0} \left| \int_{\boldsymbol{x} \in \mathbb{S}^d} \mathbf{1}_{\boldsymbol{x} \in \mathcal{S}_\alpha^{(0)}} \mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{x} \in \mathbb{S}^d} \mathbf{1}_{\boldsymbol{x} \in \mathcal{S}_\alpha^{(1)}} \mathrm{d}\boldsymbol{x} \right|.$$

In order to find an initial point inside the hyperparameters space which is close to the optimal solution, one could rely on the following heuristic argument. Assume that each individual occupancy set are disjoints.

$$\mathcal{A}_\alpha(\kappa_0) n_0 = \mathcal{A}_\alpha(\kappa_1) n_1 \quad \text{and} \quad \mathcal{A}_\alpha(\kappa_0) n_0 + \mathcal{A}_\alpha(\kappa_1) n_1 = \frac{2\pi^{d/2}}{\Gamma(d/2)}. \tag{9}$$

This gives

$$\mathcal{A}_\alpha(\kappa_0) = \frac{1}{n_0} \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad \text{and} \quad \mathcal{A}_\alpha(\kappa_1) = \frac{1}{n_1} \frac{2\pi^{d/2}}{\Gamma(d/2)}. \tag{10}$$

Therefore, κ_0 and κ_1 satisfy

$$I_{\sin^2(\theta_\alpha(\kappa_0))} \left(\frac{d-1}{2}, \frac{1}{2}\right) = \frac{2}{n_0} \quad \text{and} \quad I_{\sin^2(\theta_\alpha(\kappa_1))} \left(\frac{d-1}{2}, \frac{1}{2}\right) = \frac{2}{n_1}. \tag{11}$$

This can be numerically inverted with the `betaincinv` function of the scipy library. In our case for the MS1MV3 dataset, we have $n_0 = 27612$ and $n_1 = 64984$ and we obtain $\theta_\alpha(\kappa_0) = 1.3956$ and $\theta_\alpha(\kappa_1) = 1.3869$. From these expressions, we need to deduce values for κ_0 and κ_1. One can rely on an estimation of the left-hand side of Equation 8 as a function of κ. We compute this with a Monte-Carlo method and obtain Figure 9.

Each horizontal line of the plot intersects two hyperparameters κ_0 (for females) and κ_1 (for males) which should have good properties with respect to our problem (improving fairness while maintaining high performance). Notice that our previously chosen point $(\kappa_0 = 20, \kappa_1 = 25)$ corresponds to a mass approximately equal to 0.999, which somehow confirms that our heuristic is relevant. An extensive study of the points obtained this way would be very interesting for future works.

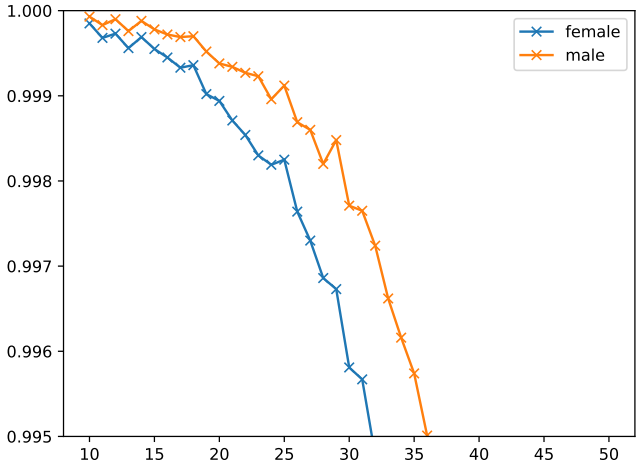


Figure 9: Approximation of the functions $\kappa \mapsto \int_{\boldsymbol{x}\in\mathbb{S}^d} V_d(\boldsymbol{x};\boldsymbol{\mu},\kappa)\mathbf{1}_{\langle \boldsymbol{x},\boldsymbol{\mu}\rangle \leq \theta_\alpha(\kappa_0)}\mathrm{d}\boldsymbol{x}$ (in blue) and $\kappa \mapsto \int_{\boldsymbol{x}\in\mathbb{S}^d} V_d(\boldsymbol{x};\boldsymbol{\mu},\kappa)\mathbf{1}_{\langle \boldsymbol{x},\boldsymbol{\mu}\rangle \leq \theta_\alpha(\kappa_1)}\mathrm{d}\boldsymbol{x}$ (in orange).

A.3.2 Trends in Figure 4

Recall that the vMF parametric interpretation of the model is that each identity is associated with a gaussian on the sphere with fixed mean and fixed concentration parameter. The images of a dataset are then seen as i.i.d. realization of the mixture of these gaussians and the loss consists in maximizing the log-likelihood. In order to control the representation power of male and female, we fix a concentration parameter κ_0 (resp. κ_1) for all male (resp. female). In Figure 4, we observe that the different metrics exhibit smooth behavior with respect to these hyperparameters. Let us give some insights on these phenomenons. In general, female are discriminated so that the maximum is realized for FAR_1: we will always place ourselves in this situation for the following heuristic reasoning, meaning that we will always assume that

$$\max(\mathrm{FAR}_0(t), \mathrm{FAR}_1(t)) = \mathrm{FAR}_1(t). \tag{12}$$

Therefore, our heuristic will not take into account the observed empirical fact that, for some specific choices of hyperparmeters, male are discriminated against. We think one could push further the reasoning to include this case but restrict the scope of our explanations in order to focus on the underlying mechanisms of the vMF loss.

Restriction to the study of $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$. We claim it is sufficient to focus on the evolution of $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$, from which the evolution of $\mathrm{FRR}_1(t)/\mathrm{FRR}_0(t)$ can be deduced, at least at the heuristic level developed here. Two cases may occur:

- If $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$ increases, it means that there are more False Acceptance among females. From a geometric viewpoint, this means that females are more spread around than males. Therefore, there will be more False Reject among males who are more concentrated. Thus, when $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$ increases, $\mathrm{FRR}_1(t)/\mathrm{FRR}_0(t)$ decreases.
- If $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$ decreases, it means that there are less False Acceptance among females. From a geometric viewpoint, this means that females are more concentrated than males. Therefore, there will be less False Reject among males who are less concentrated. Thus, when $\mathrm{FAR}_1(t)/\mathrm{FAR}_0(t)$ decreases, $\mathrm{FRR}_1(t)/\mathrm{FRR}_0(t)$ increases.

These two observations are confirmed by the graphical representations of Figure 4.

Suppose that κ_1 is increased by a small amount $\Delta\kappa_1$ while κ_0 remains fixed.
We will denote by $\mathrm{FAR}_a^{\kappa_1}$ the False Acceptance Rate curve of subgroup a for the hyperparameters

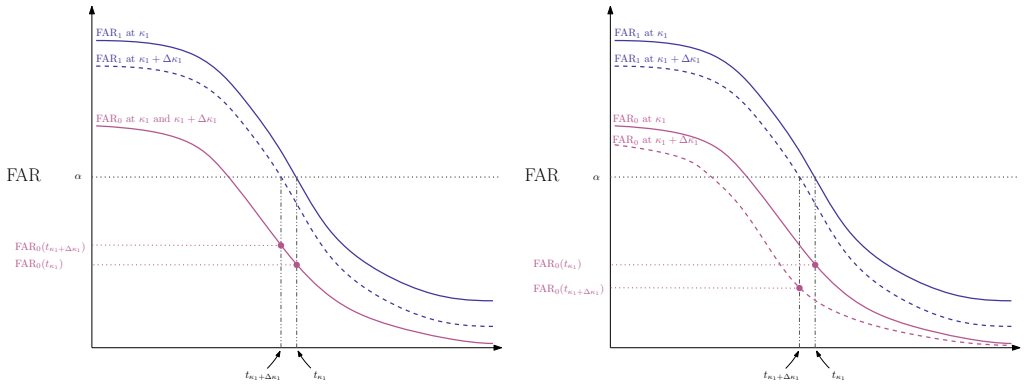
choice (κ_0, κ_1) and by $\mathrm{FAR}_a^{\kappa_1+\Delta\kappa_1}$ the False Acceptance Rate curve of subgroup a for the hyperparameters choice $(\kappa_0, \kappa_1 + \Delta\kappa_1)$.
The representation with hyperparameters $(\kappa_0, \kappa_1 + \Delta\kappa_1)$ increases the concentration parameter of females. As a result, the images stemming from a same female identity should be closer from one another, leading to a better FAR performance. Therefore, one should have:

$$\forall t \in [0,1], \quad \mathrm{FAR}_1^{\kappa_1+\Delta\kappa_1}(t) < \mathrm{FAR}_1^{\kappa_1}(t). \tag{13}$$

Let us denote by t_{κ_1} and $t_{\kappa_1+\Delta\kappa_1}$ the points satisfying:

$$\mathrm{FAR}_1^{\kappa_1}(t_{\kappa_1}) = \alpha \quad \text{and} \quad \mathrm{FAR}_1^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1}) = \alpha.$$

Using Equation 12 and Equation 13, this implies that $t_{\kappa_1+\Delta\kappa_1} < t_{\kappa_1}$, as illustrated in Figure 10.



(a) When κ_1 is small, the FAR_0 curve is not perturbed.

(b) When κ_1 is large enough, the FAR_0 curve starts to be perturbed.

Figure 10: Heuristic explanation of the FARs evolution at fixed κ_0 and when κ_1 increases.

We can now distinguish two situations depending on the magnitude of κ_1.

- If κ_1 is small, its variation does not affect the representation of males at least at a first order approximation. In that case $\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1}) = \mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})$. Since $\mathrm{FAR}_0^{\kappa_1}$ is nonincreasing, we deduce that $\mathrm{FAR}_0(t_{\kappa_1+\Delta\kappa_1}) > \mathrm{FAR}_0(t_{\kappa_1})$, which finally implies that:

$$\frac{\mathrm{FAR}_1^{\kappa_1}(t_{\kappa_1})}{\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1})} = \frac{\alpha}{\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1})} > \frac{\alpha}{\mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})} = \frac{\mathrm{FAR}_1^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})}{\mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})}.$$

- If κ_1 is large enough, tightening the representations of females among themselves starts to affect the males representation. Indeed, they enjoy more space and can therefore be better spread, which implies that $\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1}) > \mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})$, as illustrated in Figure 10 (b). As a result:

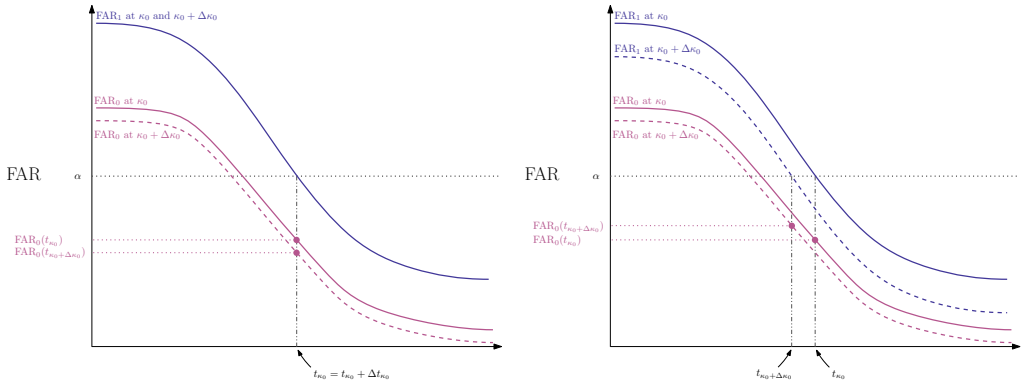
$$\frac{\mathrm{FAR}_1^{\kappa_1}(t_{\kappa_1})}{\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1})} = \frac{\alpha}{\mathrm{FAR}_0^{\kappa_1}(t_{\kappa_1})} < \frac{\alpha}{\mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})} = \frac{\mathrm{FAR}_1^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})}{\mathrm{FAR}_0^{\kappa_1+\Delta\kappa_1}(t_{\kappa_1+\Delta\kappa_1})}.$$

The two previous points are confirmed by the top left corner graphical representation of Figure 4: for all fixed values of κ_0, the curves start by decreasing when κ_1 increases, then begin an increasing phase when κ_1 becomes sufficiently large.

Suppose that κ_0 is increased by a small amount $\Delta\kappa_0$ while κ_0 remains fixed.
We will denote by $\mathrm{FAR}_a^{\kappa_0}$ the False Acceptance Rate curve of subgroup a for the hyperparameters choice (κ_0, κ_1) and by $\mathrm{FAR}_a^{\kappa_0+\Delta\kappa_0}$ the False Acceptance Rate curve of subgroup a for the hyperparameters choice $(\kappa_0 + \kappa_0, \kappa_1)$.
The representation with hyperparameters $(\kappa_0 + \Delta\kappa_0, \kappa_1)$ increases the concentration parameter of males. As a result, the images stemming from a same male identity should be closer from one another, leading to a better FAR performance. Therefore, one should have:

$$\forall t \in [0,1], \quad \mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t) < \mathrm{FAR}_0^{\kappa_0}(t). \tag{14}$$

As before, we can distinguish two situations depending on the magnitude of κ_0.



(a) When κ_1 does not increase to much, the FAR_0 curve is not perturbed.

(b) When κ_1 is large enough, the FAR_0 curve starts to be perturbed.

Figure 11: Heuristic explanation of the FARs evolution at fixed κ_0 and when κ_1 increases.

- If κ_0 is small, one can suppose that females are not affected by its variation, meaning that $\mathrm{FAR}_1^{\kappa_0} = \mathrm{FAR}_1^{\kappa_0+\Delta\kappa_0}$ at a first order approximation (see (a) of Figure 11 for an illustration). In that case, $t_{\kappa_0} = t_{\kappa_0+\Delta\kappa_0}$, and Equation 14 implies that $\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0}) < \mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})$. As a result:

$$\frac{\mathrm{FAR}_1^{\kappa_0}(t_{\kappa_0})}{\mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})} = \frac{\alpha}{\mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})} < \frac{\alpha}{\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})} = \frac{\mathrm{FAR}_1^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})}{\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})}.$$

- If κ_0 is large enough, tightening the representations of males among themselves starts to affect the females representation: they have more space to spread around (see (b) of Figure 11). As a result, one can have $\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0}) > \mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})$

$$\frac{\mathrm{FAR}_1^{\kappa_0}(t_{\kappa_0})}{\mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})} = \frac{\alpha}{\mathrm{FAR}_0^{\kappa_0}(t_{\kappa_0})} > \frac{\alpha}{\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})} = \frac{\mathrm{FAR}_1^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})}{\mathrm{FAR}_0^{\kappa_0+\Delta\kappa_0}(t_{\kappa_0+\Delta\kappa_0})}.$$

A.4 Fairness evaluation on LFW

Once some relevant pairs (κ_0, κ_1) are chosen for ArcFace ResNet100 using the dataset IJB-C, we evaluate them on the LFW dataset Huang et al. (2008) which contains ground-truth gender labels. The official LFW protocol only considers a few matching pairs among all the possible pairs given the whole LFW dataset. The number of female images is typically not enough to get good estimates of our fairness metrics. To overcome this, we consider all possible same-gender matching pairs among the whole LFW dataset. Doing so, we obtain 9.8k female genuine pairs, 232k male genuine pairs, 4.4M female impostor pairs and 52M male impostor pairs.

A.5 Additional numerical results

In this section, we provide more numerical experiments, varying the evaluation dataset (LFW, IJB-C, IJB-B) and different kinds of pre-trained models (ArcFace with several ResNet architectures, other pre-trained models with MobileFaceNet backbone).

FAR level:		10^{-4}			10^{-3}		
model		FRR@FAR (%)	BFRR	BFAR	FRR@FAR (%)	BFRR	BFAR
ArcFace	original	**3.94**	1.97	4.06	**2.68**	1.79	2.04
	(15,20)	4.90	2.33	**1.17**	2.90	1.98	1.15
	(25,20)	4.34	**1.62**	11.86	2.76	**1.60**	5.58
	(45,30)	5.20	1.92	1.25	3.53	1.91	**1.07**
AdaCos	original	18.85	1.18	5.44	**9.74**	1.24	3.84
	(15,20)	20.75	1.30	**2.06**	10.31	1.42	2.20
	(25,20)	20.28	**1.02**	13.00	10.09	**1.06**	7.80
	(45,30)	**17.48**	1.28	2.86	9.85	1.33	**2.06**
CosFace	original	**15.67**	1.24	3.08	**8.55**	1.35	2.54
	(15,20)	19.52	1.35	**2.75**	10.24	1.41	**2.33**
	(25,20)	20.57	**1.03**	86.69	10.24	**1.04**	13.61
	(45,30)	17.27	1.12	8.67	9.69	1.11	4.29
Curricular	original	**17.69**	1.19	8.18	**9.26**	1.30	4.21
	(15,20)	19.97	1.33	**3.23**	10.37	1.42	**2.23**
	(25,20)	20.35	**1.04**	20.88	10.02	**1.03**	9.54
	(45,30)	18.07	1.18	5.29	9.99	1.22	3.33

Table 3: IJBC 1:1 protocol for ArcFace with ResNet100 backbone and different pre-trained models with MobileFaceNet backbone. By "original" we mean no Ethical Module is added to the pre-trained model. The tuples correspond to the choices of κ_0 (first argument) and κ_1 (second argument).

FAR level:		10^{-4}			10^{-3}		
model		FRR@FAR (%)	BFRR	BFAR	FRR@FAR (%)	BFRR	BFAR
R100	original	**0.063**	10.76	3.98	**0.052**	2.23	1.81
	(15,20)	0.119	12.73	**1.72**	0.067	8.43	**1.04**
	(25,20)	0.076	**5.35**	29.33	0.052	**1.94**	3.96
	(45,30)	0.129	13.47	2.99	0.067	6.02	1.24
R50	original	**0.078**	10.27	4.72	0.059	4.17	1.81
	(15,20)	0.151	11.22	**2.11**	0.072	9.16	**1.19**
	(25,20)	0.100	**5.89**	33.65	**0.058**	**4.11**	5.24
	(45,30)	0.164	9.18	2.44	0.081	5.15	1.20
R34	original	**0.104**	11.81	7.62	**0.063**	8.64	2.17
	(15,20)	0.204	14.27	**3.31**	0.087	17.56	1.59
	(25,20)	0.163	**5.63**	43.55	0.069	**8.09**	6.43
	(45,30)	0.226	8.85	4.42	0.095	8.80	**1.02**
R18	original	**0.214**	11.16	2.80	**0.116**	7.53	1.93
	(15,20)	0.465	11.15	**1.59**	0.197	10.60	**1.34**
	(25,20)	0.310	**4.44**	24.59	0.125	**6.53**	7.57
	(45,30)	0.349	6.69	4.21	0.162	6.92	1.76

Table 4: Evaluation on LFW for ArcFace on different ResNet architectures. By "original" we mean no Ethical Module is added to the pre-trained model. The tuples correspond to the choices of κ_0 (first argument) and κ_1 (second argument).

FAR level:		10^{-4}			10^{-3}		
model		FRR@FAR (%)	BFRR	BFAR	FRR@FAR (%)	BFRR	BFAR
R100	original	**3.94**	1.97	4.06	**2.68**	1.79	2.04
	(15,20)	4.90	2.33	**1.17**	2.90	1.98	1.15
	(25,20)	4.34	**1.62**	11.86	2.76	**1.60**	5.58
	(45,30)	5.20	1.92	1.25	3.53	1.91	**1.07**
R50	original	**4.29**	1.81	3.41	**3.00**	1.88	1.95
	(15,20)	5.56	2.18	1.28	3.40	2.18	**1.00**
	(25,20)	4.91	**1.49**	10.87	3.19	**1.50**	6.49
	(45,30)	5.41	1.73	**1.24**	3.71	1.77	1.09
R34	original	**4.95**	1.72	2.83	**3.47**	1.77	1.88
	(15,20)	6.38	2.05	**1.17**	3.85	2.04	**1.06**
	(25,20)	5.67	**1.45**	13.69	3.60	**1.50**	5.86
	(45,30)	6.13	1.62	1.70	4.24	1.69	**1.06**
R18	original	**6.64**	1.68	3.81	**4.41**	1.58	2.37
	(15,20)	8.64	1.83	**1.39**	4.96	1.88	**1.43**
	(25,20)	8.27	**1.19**	16.25	4.76	**1.26**	10.94
	(45,30)	7.46	1.50	3.16	4.97	1.56	1.85

Table 5: IJBC 1:1 protocol for ArcFace on different ResNet architectures. By "original" we mean no Ethical Module is added to the pre-trained model. The tuples correspond to the choices of κ_0 (first argument) and κ_1 (second argument).

		FRR@FAR (%)	
FAR level:		10^{-4}	10^{-3}
R100	original	5.38	3.78
	(15,20)	6.79	4.11
	(25,20)	6.00	3.84
	(45,30)	7.03	4.81
R50	original	5.95	4.20
	(15,20)	7.58	4.71
	(25,20)	6.71	4.26
	(45,30)	7.34	5.10
R34	original	6.72	4.63
	(15,20)	8.54	5.18
	(25,20)	7.62	4.60
	(45,30)	8.11	5.57
R18	original	8.59	5.76
	(15,20)	11.12	6.53
	(25,20)	10.94	6.01
	(45,30)	9.72	6.35

Table 6: IJB-B 1:1 protocol for ArcFace on different ResNet architectures. By "original" we mean no Ethical Module is added to the pre-trained model. The tuples correspond to the choices of κ_0 (first argument) and κ_1 (second argument).

		FRR@FAR (%)	
FAR level:		10^{-4}	10^{-3}
ArcFace	original	5.38	3.78
	(15,20)	6.79	4.11
	(25,20)	6.00	3.84
	(45,30)	7.03	4.81
AdaCos	original	22.98	12.27
	(15,20)	24.06	12.78
	(25,20)	24.41	12.78
	(45,30)	21.25	12.44
CosFace	original	18.85	10.65
	(15,20)	23.38	12.51
	(25,20)	26.10	13.01
	(45,30)	21.22	12.27
Curricular	original	12.20	11.42
	(15,20)	24.50	12.56
	(25,20)	24.91	12.35
	(45,30)	21.88	11.97

Table 7: IJB-B 1:1 protocol for ArcFace with ResNet100 backbone and different pre-trained models with MobileFaceNet backbone. By "original" we mean no Ethical Module is added to the pre-trained model. The tuples correspond to the choices of κ_0 (first argument) and κ_1 (second argument).