- ▹ **Blur ($\sqrt{\varepsilon}$).** Best results are achieved with `blur` ≈ 0.003. Larger values lead to more bias, smaller values cause more variance, so F1 decreases in either direction.

- ▹ **Orthogonality weight.** Varying λ shifts macro F1 by $\approx < 0.003$. This implies that the OT scalar already explains the signal.

- ▹ **Style dimension.** Performance is effectively flat between 64 and 256 dimensions, which means that stylistic information is low-rank.

- ▹ **Projection dimension.** There is no gain beyond $n_\pi = 32$, and larger values simply lead to slower training.

Apart from `blur` tuning in the feature-free balanced OT method, OT variants are not particularly sensitive to orthogonality weight/projection dimension/style dimension tuning in the context of style shift detection.