Fairness Metrics vs Flip Percentage (Bias-Bias)

Independence (M_ind) vs Flip Percentage

Before: 0.2948 ± 0.0193
After: 0.1120 ± 0.0092
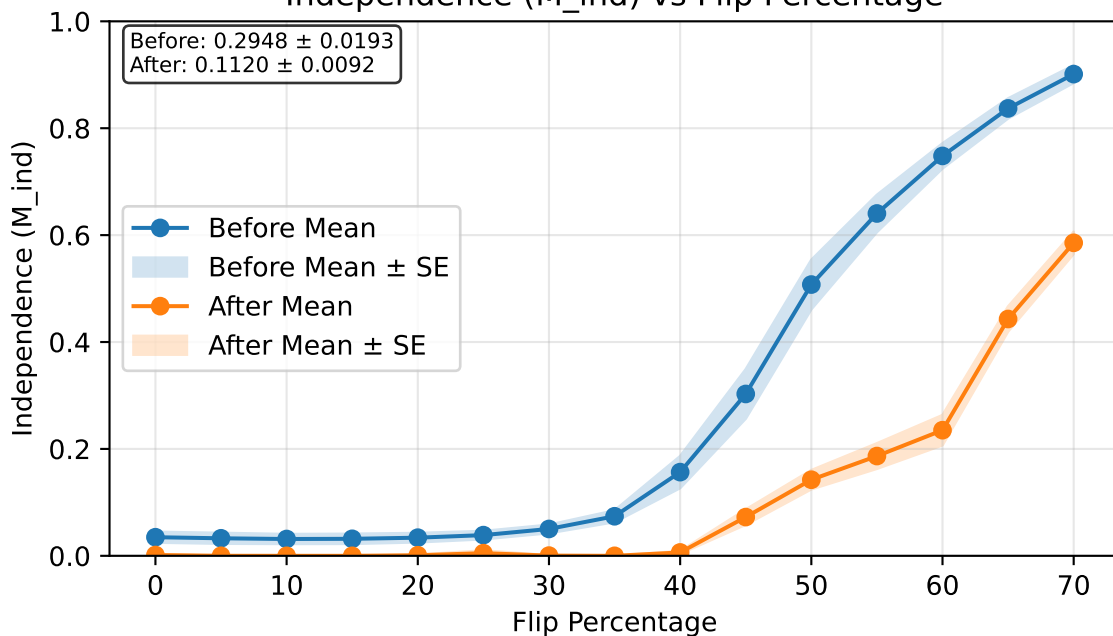
Separation (M_sep) vs Flip Percentage

Before: 0.5770 ± 0.0420
After: 0.1637 ± 0.0155
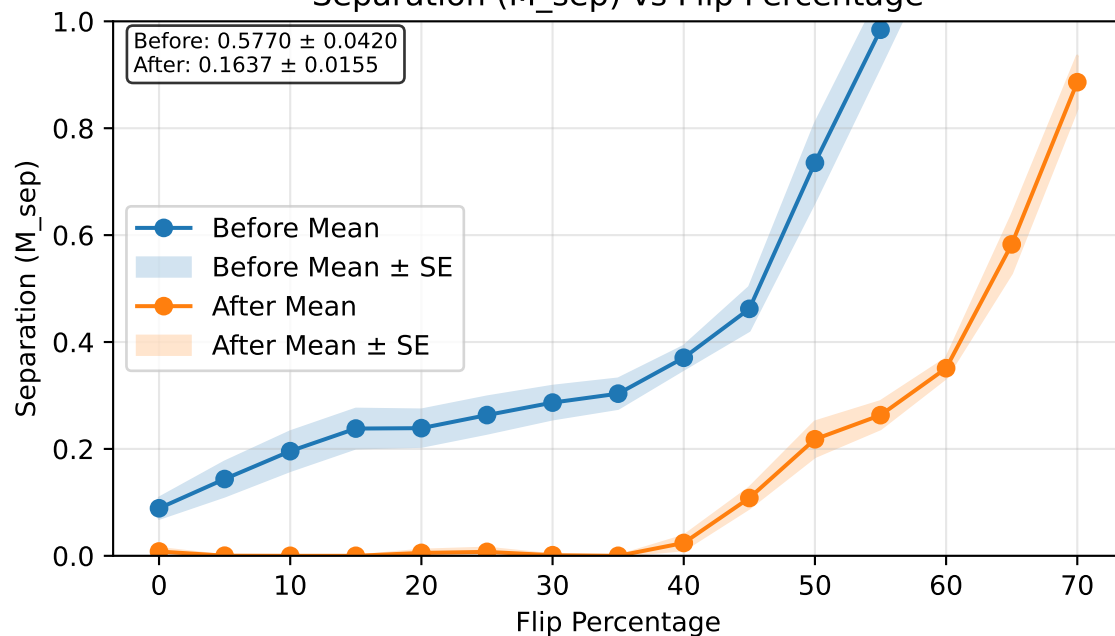
Sufficiency (M_suff) vs Flip Percentage

Before: 0.4743 ± 0.0364
After: 0.4150 ± 0.0492