By Jensen’s Inequality: , so regardless of the correlation, we have , with equality only when the number of matched samples is very large (I think). The observed test statistic therefore has more weight in the tails and for a given critical value , , with equality only in the limit (with m).

By using the 20th quantile estimator, we set . Equivalently, we set . In other words, we force our inflation factor to be conservative, allowing an “inflated” estimate only 20% of the time.

There may be more room to show how exactly works probabilistically in controlling the Type I error rate, but I’m not sure we need to go too far in that direction.