Stability of Measure Comparisons for wig+100 on nfcorpus (p=0.25 Count Delta 0.5 600 1200 0.4 1800 Swap Probability 2400 0.3 3000 0.2 0.1 0.0 0.2 0.4 0.6 8.0 1.0 1.2

Average Performance Delta